

Protein Sequence Searches - 10/8/04

All of the sequence databases on the ABSS have been updated. A change has occurred in the protein databases.

- Two protein databases, SPTREMBL and SwissProt, are now produced as a single, merged database called UniProt.
- Results from UniProt have the file extension **.rup**.
- Sequences in UniProt are identified by the same ID that had been used in SPTREMBL or SwissProt.
- In instances where the database curators have determined that an SPTREMBL record and a SwissProt record represent the same sequence, the two records have been merged into one. Both IDs are present in the record. Any differences found between the two sequences are recorded in the FT (feature table) fields.

If you have any questions regarding these changes or your results, please contact any STIC searcher.

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 22, 2004, 18:22:47 ; Search time 196 Seconds

(without alignments)
1441.372 Million cell updates/sec

Title: US-10-017-724-6

Perfect score: 2676

Sequence: 1 MKRMVSWSFHLKTMKHL.....GSWYSMRKMKIRPFPPQ 491

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt_02.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2671	99.8	491	1 FIBB_HUMAN	P02675 homo sapien
2	2259	84.4	481	2 Q8K0E8	Q8K0E8 mus musculus
3	2176	81.3	479	1 FIBB_RAT	P14480 rattus norv
4	2135	79.8	468	1 FIBB_BOVIN	P02676 bos taurus
5	2080	77.7	503	2 Q7TME5	Q7TME5 rattus norv
6	1744	65.2	463	1 FIBB_CHICK	Q02020 gallus gall
7	1734.5	64.8	488	2 Q91589	Q91589 xenopus lae
8	1464	54.7	485	2 Q6NYE1	Q6NYE1 brachydanio
9	1464	54.7	485	2 A4H66629	A4H66629 brachydan
10	1384.5	51.7	321	2 Q7SYN8	Q7SYN8 xenopus lae
11	1279	47.8	236	2 Q912P1	Q912P1 mus musculus
12	1211	45.3	477	1 FIBB_PETMA	P02678 petromyzon
13	657	24.6	431	2 Q7ZVGT	Q7ZVGT brachydanio
14	656.5	24.5	435	2 Q93568	Q93568 gallus gall
15	631.5	23.6	436	2 Q8VCM7	Q8VCM7 mus musculus
16	630.5	23.6	437	2 A4P35744	A4P35744 homo sapi
17	630.5	23.6	453	1 FIBB_HUMAN	P02679 homo sapien
18	611	22.8	439	2 Q7SZ13	Q7SZ13 xenopus lae
19	609.5	22.8	399	2 Q7Z664	Q7Z664 homo sapien
20	609	22.8	444	1 FIBG_BOVIN	P12799 bos taurus
21	607.5	22.7	438	1 FIBG_XENLA	P17634 xenopus lae
22	601.5	22.5	445	1 FIBG_RAT	P02680 rattus norv
23	598.5	22.4	432	1 FIBG_PETMA	P04115 petromyzon
24	575	21.5	316	2 Q28529	Q28529 mustela put
25	563	21.0	313	2 Q6G1L5	Q6G1L5 xenopus lae
26	545	20.4	760	2 Q7ZTR1	Q7ZTR1 xenopus lae
27	544.5	20.3	866	1 FIBA_HUMAN	P02671 homo sapien
28	535	20.0	764	2 Q6GNK6	Q6GNK6 xenopus lae
29	534.5	20.0	735	2 Q28763	Q28763 papio cynoc
30	531.5	19.9	493	1 ANL2_MOUSE	Q9r045 mus musculus
31	531.5	19.9	493	2 Q8BM09	Q8bm09 mus musculus

32	530.5	19.8	493	1 ANL2_HUMAN	Q9uku9 homo sapien
33	529.5	19.8	493	2 Q9JU03	Q9jj03 rattus norv
34	523.5	19.6	782	2 Q7TQ70	Q7tq70 rattus norv
35	522	19.5	314	2 Q71KU9	Q71ku9 mus musculus
36	522	19.5	314	2 AAQ05798	AaQ05798 mus muscu
37	517	19.3	314	2 Q8VC25	Q8vc25 mus musculu
38	516.5	19.3	741	1 FIBA_CHICK	P14448 gallus gall
39	512.5	19.2	782	1 FIBA_RAT	P06399 rattus norv
40	512	19.1	312	1 FGLI_HUMAN	Q08830 homo sapien
41	512	19.1	312	2 Q8NG32	Q8ng32 homo sapien
42	512	19.1	312	2 AAP35281	AaP35281 homo sapi
43	510.5	19.1	491	2 Q95841	Q95841 homo sapien
44	510.5	19.1	491	2 AAQ88645	AaQ88645 homo sapi
45	505	18.9	236	2 Q8NZJ9	Q8nzj9 homo sapien

ALIGNMENTS

RESULT 1

ID	FIBB_HUMAN	STANDARD;	PRT;	491 AA.
AC	P02675;			
DT	21-JUL-1986 (Rel. 01, Created)			
DT	01-JUL-1993 (Rel. 26, Last sequence update)			
DT	01-OCT-2004 (Rel. 45, Last annotation update)			
DE	Fibrinogen beta chain precursor [Contains: Fibrinopeptide B].			
GN	Name=FGB;			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=91344740; PubMed=2102623;			
RA	Chung D.W., Harris J.E., Davie E.W.;			
RT	"Nucleotide sequences of the three genes coding for human			
RT	fibrinogen.";			
RL	Adv. Exp. Med. Biol. 281:39-48(1990).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=83283433; PubMed=6688356;			
RA	Chung D.W., Que B.G., Rixon M.W., Mace M. Jr., Davie E.W.;			
RT	"Characterization of complementary deoxyribonucleic acid and genomic			
RT	deoxyribonucleic acid for the beta chain of human fibrinogen.";			
RL	Biochemistry 22:3244-3250(1983).			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RX	Chung D.W., Harris J.E., Davie E.W.;			
RA	"Nucleotide sequences of the three genes coding for human			
RT	fibrinogen.";			
RL	(in) Liu C.Y., Chien S. (eds.);			
RL	Fibrinogen, thrombosis, coagulation and fibrinolysis, pp.39-48, Plenum			
RL	Press, New York (1991).			
RN	[4]			
RP	SEQUENCE FROM N.A., AND VARIANTS SER-100; HIS-170; LEU-265 AND			
RP	LYS-478.			
RA	Rieder M.J., Carrington D.P., Chung M.-W., Lee K.L., Yi Q.,			
RA	Nickerson D.A.;			
RT	"SeattlesNPs. NHLBI HL66682 program for genomic applications, UW-			
RT	FHCRC, Seattle, WA (URL: http://pga.gs.washington.edu).";			
RL	Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.			
RN	[5]			
RP	SEQUENCE OF 31-491, AND CARBOHYDRATE-LINKAGE SITE.			
RA	Henschen A., Lottspeich F., Southan C., Topfer-Petersen E.;			
RT	"Human fibrinogen: sequence, sulfur bridges, glycosylation and some			
RT	structural variants.";			
RL	(in) Peeters H. (eds.);			
RL	Protides of the biological fluids, Proc. 28th colloquium, pp.51-56,			
RL	Pergamon Press, Oxford (1980).			
RN	[6]			
RP	SEQUENCE OF 31-491.			
RX	MEDLINE=79124640; PubMed=420779;			

- RA Watt K.W.K., Takagi T., Doolittle R.F.;
RT "Amino acid sequence of the beta chain of human fibrinogen.";
RL Biochemistry 18:68-76(1979).
RN [7]
- RP SEQUENCE OF 31-148, AND DISULFIDE BONDS.
RX MEDLINE=76225080; PubMed=936108;
RA Blomback B., Hessel B., Hoeg D.;
RT "Disulfide bridges in NH2-terminal part of human fibrinogen.";
RL Thromb. Res. 8:639-658(1976).
RN [8]
- RP SEQUENCE OF 1-38 FROM N.A.
RX MEDLINE=871146483; PubMed=9329722;
RA Huber P., Dalmon J., Courtois G., Laurent M., Assouline Z.,
RA Marguerie G.;
RT "Characterization of the 5'-flanking region for the human fibrinogen
beta gene.";
RL Nucleic Acids Res. 15:1615-1625(1987).
RN [9]
- RP SEQUENCE OF 31-44.
RA Blomback B., Blomback M., Grondahl N.J., Guthrie C., Hinton M.;
RT "Studies on fibrinopeptides from primates.";
RL Acta Chem. Scand. 19:1788-1789(1965).
RN [10]
- RP REVIEW, AND DISULFIDE BONDS.
RX MEDLINE=83254370; PubMed=6575689;
RA Henschen A., Lottspeich F., Kehl M., Southan C.;
RT "Covalent structure of fibrinogen.";
RL Ann. N. Y. Acad. Sci. 408:28-43(1983).
RN [11]
- RP DISULFIDE BONDS.
RX MEDLINE=77245999; PubMed=891553;
RA Gaardlund B., Hessel B., Marguerie G., Murano G., Blomback B.;
RT "Primary structure of human fibrinogen. Characterization of disulfide-
containing cyanogen-bromide fragments.";
RL Eur. J. Biochem. 77:595-610(1977).
RN [12]
- RP DISULFIDE BONDS.
RA Doolittle R.F., Takagi T., Watt K.W.K., Bouma H. III, Cottrell B.A.,
RA Cassman K.G., Goldbaum D.M., Doolittle L.R., Friesner S.J.;
RT "The structures of fibrinogen and fibrin.";
RL (in) Magnusson S., Ottesen M., Folmann B., Dano K., Neurath H.
(eds.);
RL Regulatory proteolytic enzymes and their inhibitors, pp.163-172,
RL Pergamon Press, New York (1978).
RN [13]
- RP REVIEW, EM STRUCTURE, POLYMERIZATION, AND LIGANDS.
RX MEDLINE=84305751; PubMed=6383194;
RA Doolittle R.F.;
RT "Fibrinogen and fibrin.";
RL Annu. Rev. Biochem. 53:195-229(1984).
RN [14]
- RP X-RAY CRYSTALLOGRAPHY (2.9 ANGSTROMS) OF 164-491.
RX MEDLINE=97472408; PubMed=9333233;
RA Spraggon G., Everse S.J., Doolittle R.F.;
RT "Crystal structures of fragment D from human fibrinogen and its
crosslinked counterpart from fibrin.";
RL Nature 389:455-462(1997).
RN [15]
- RP X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS) OF 164-491.
RX MEDLINE=98292395; PubMed=9628725;
RA Everse S.J., Spraggon G., Veerapandian L., Riley M., Doolittle R.F.;
RT "Crystal structure of fragment double-D from human fibrin with two
different bound ligands.";
RL Biochemistry 37:8637-8642(1998).
RN [16]
- RP X-RAY CRYSTALLOGRAPHY.
RX MEDLINE=99175089; PubMed=10074346;
RA Everse S.J., Spraggon G., Veerapandian L., Doolittle R.F.;
RT "Conformational changes in fragments D and double-D from human
fibrin(ogen) upon binding the peptide ligand Gly-His-Arg-Pro-amide.";
RL Biochemistry 38:2941-2946(1999).
RN [17]
- RP INTERACTION WITH FBLN1.
RX MEDLINE=95370284; PubMed=7642629;
RA Tran H., Tanaka A., Litvinovich S.V., Medved L.V., Haudenschild C.C.,
RA Agraves W.S.;
RT "The interaction of fibulin-1 with fibrinogen. A potential role in
hemostasis and thrombosis.";
RL J. Biol. Chem. 270:19458-19464(1995).
RN [18]
- RP VARIANT BALTIMORE-2 LYS-478.
RX MEDLINE=89058942; PubMed=3194892;
RA Schmeller C.H., Ebert R.F., Bell W.R.;
RT "A polymorphism at B beta 448 of fibrinogen identified during
structural studies of fibrinogen Baltimore II.";
RL Thromb. Res. 52:173-177(1988).
RN [19]
- RP VARIANT ISE ARG-45.
RX MEDLINE=91208409; PubMed=2018836;
RA Yoshida N., Wada H., Morita K., Hirata H., Matsuda M., Yamazumi K.,
RA Asakura S., Shirakawa S.;
RT "A new congenital abnormal fibrinogen Ise characterized by the
replacement of B beta glycine-15 by cysteine.";
RL Blood 77:1958-1963(1991).
RN [20]
- RP VARIANT NAPLES THR-98.
RX MEDLINE=92340664; PubMed=1634610;
RA Koopman J., Haverkate F., Lord S.T., Grimbergen J., Mannucci P.M.;
RT "Molecular basis of fibrinogen Naples associated with defective
thrombin binding and thrombophilia. Homozygous substitution of B beta
68 Ala-->Thr.";
RL J. Clin. Invest. 90:238-244(1992).
RN [21]
- RP VARIANTS IJMUUDEN CYS-44 AND NIJMEGEN CYS-74.
RX MEDLINE=92228809; PubMed=1565641;
RA Koopman J., Haverkate F., Grimbergen J., Engesser L., Novakova I.,
RA Kerst A.F.J.A., Lord S.T.;
RT "Abnormal fibrinogens Ijmuuden (B beta Arg14-->Cys) and Nijmegen (B
beta Arg44-->Cys) form disulfide-linked fibrinogen-albumin
complexes.";
RL Proc. Natl. Acad. Sci. U.S.A. 89:3478-3482(1992).
RN [22]
- RP VARIANT NEW YORK-1 39-GLY--LEU-102 DEL.
RX MEDLINE=85157605; PubMed=3156856;
RA Liu C.Y., Koehn J.A., Morgan F.J.;
RT "Characterization of fibrinogen New York 1. A dysfunctional fibrinogen
with a deletion of B beta(9-72) corresponding exactly to exon 2 of the
gene.";
RL J. Biol. Chem. 260:4390-4396(1985).
RN [23]
- RP VARIANTS GLU-2; LEU-265 AND LYS-478.
RX MEDLINE=99318093; PubMed=10391209; DOI=10.1038/10290;
RA Cargill M., Altshuler D., Ireland J., Sklar P., Ardlie K., Patil N.,
RA Shaw N., Lane C.R., Lim E.P., Kalyanaram N., Nemesh J., Ziaugra L.,
RA Friedland L., Rolfe A., Warrington J., Lipshutz R., Daley G.Q.,
RA Lander E.S.;
RT "Characterization of single-nucleotide polymorphisms in coding regions
of human genes.";
RL Nat. Genet. 22:231-238(1999).
RN [24]
- RP ERRATUM.
RX MEDLINE=20129589; PubMed=10666208;
RA Duga S., Asselta R., Santagostino E., Zeinali S., Simonc T.,
RA Malcovati M., Mannucci P.M., Tenchini M.L.;
RT "Missense mutations in the human beta fibrinogen gene cause congenital
afibrinogenemia by impairing fibrinogen secretion.";
RL Blood 95:1336-1341(2000).
RN [26]
- RP VARIANT CONGENITAL AFIBRINOGENEMIA CYS-196.
RX MEDLINE=20129589; PubMed=10666208;
RA Duga S., Asselta R., Santagostino E., Zeinali S., Simonc T.,
RA Malcovati M., Mannucci P.M., Tenchini M.L.;
RT "Missense mutations in the human beta fibrinogen gene cause congenital
afibrinogenemia by impairing fibrinogen secretion.";
RL Blood 95:1336-1341(2000).
RN [26]

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 22, 2004, 18:25:22 ; Search time 42 Seconds

(without alignments)
1124.820 Million cell updates/sec

Title: US-10-017-724-6

Perfect score: 2676

Sequence: 1 MKRWVSWSFHLKTKMKHLL.....GWSYMKRWKMKIRPFPPQ 491

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR 79:*

1: PIR1:*

2: PIR2:*

3: PIR3:*

4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2671	99.8	491	1 FGHUB	fibrinogen beta ch
2	2132	79.7	468	1 FGHUB	fibrinogen beta ch
3	1744	65.2	463	2 A38463	fibrinogen beta ch
4	1546	57.8	328	2 A05299	fibrinogen beta ch
5	1218	45.5	479	2 A25052	fibrinogen beta ch
6	630.5	23.6	437	1 FGHUG	fibrinogen gamma-A
7	630.5	23.6	453	1 FGHUGB	fibrinogen gamma-B
8	609	22.8	444	2 S05313	fibrinogen gamma-B
9	607.5	22.7	438	2 A32670	fibrinogen gamma c
10	598.5	22.4	432	1 FGLMGS	fibrinogen gamma c
11	590.5	22.1	437	1 FGRTGA	fibrinogen gamma-A
12	590.5	22.1	445	1 FGRTGB	fibrinogen gamma-B
13	544.5	20.3	866	2 D4234	fibrinogen alpha c
14	529.5	19.8	493	2 JC8027	type 1 angiotensin
15	532	19.1	312	2 JN0596	fibrinogen-related
16	496	18.5	641	1 A41932	fibrinogen alpha-I
17	464	17.3	432	2 A27447	cytotoxic T-lympho
18	461	17.2	439	2 I37391	fibrinogen-like pr
19	460	17.2	432	2 I56934	fibrinogen-like pr
20	437	16.3	334	2 JCS980	ficollin-A precursor
21	430.5	16.1	326	2 S61517	ficollin-1 precursor
22	429.5	16.1	333	2 A47172	transforming growt
23	426.5	15.9	326	2 B47172	ficollin-beta - pig
24	422.5	15.8	1353	1 JH0675	restrictin precursor
25	414	15.5	1356	2 A45445	janusin precursor
26	411.5	15.4	2019	1 JQ1322	tenascin precursor
27	406.5	15.2	1810	1 A32230	tenascin precursor
28	401	15.0	91	2 S69115	fibrinogen beta ch
29	395.5	14.8	1746	1 S19694	tenascin precursor

ALIGNMENTS

RESULT 1

FGHUB

fibrinogen beta chain precursor [validated] - human

N;Alternate names: coagulation factor I

C;Contains: fibrinopeptide B

C;Species: Homo sapiens (man)

C;Date: 24-Apr-1984 #sequence revision 31-Mar-1993 #text change 09-Jul-2004

C;Accession: B43568; A90469; I37389; A94433; A90437; A94309; G54223; A03121; B371

R;Chung, D.W.; Harris, J.E.; Davie, E.W.

Adv. Exp. Med. Biol. 281, 39-48, 1990

A;Title: Nucleotide sequences of the three genes coding for human fibrinogen.

A;Reference number: A43568; MUID:91344740; PMID:2102623

A;Accession: B43568

A;Molecule type: DNA

A;Residues: 9-191,'A',193-491 <CHU>

A;Cross-references: UNIPROT:P02675

R;Chung, D.W.; Que, B.G.; Rixon, M.W.; Mace Jr., M.; Davie, E.W.

Biochemistry 22, 3244-3250, 1983

A;Title: Characterization of complementary deoxyribonucleic acid and genomic deoxyribonu

A;Reference number: A90469; MUID:83283433; PMID:6688356

A;Accession: A90469

A;Molecule type: DNA

A;Residues: 1-38 <CH1>

A;Accession: B90469

A;Molecule type: mRNA

A;Residues: 9-191,'A',193-491 <CH2>

A;Cross-references: GB:J00129; NID:g182429; PIDN:AAA52429.1; PID:g182430

R;Huber, P.; Dalmon, J.; Courtois, G.; Laurent, M.; Assouline, Z.; Marguerie, G.

Nucleic Acids Res. 15, 1615-1625, 1987

A;Title: Characterization of the 5'-flanking region for the human fibrinogen beta gene.

A;Reference number: I37389; MUID:87146483; PMID:3029722

A;Accession: I37389

A;Status: translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-38 <HUB>

A;Cross-references: EMBL:X05018; NID:g31400; PIDN:CAA28674.1; PID:g31401

R;Henschen, A.; Lottspeich, F.; Southan, C.; Tøpfer-Petersen, E.

In Protides of the Biological Fluids, Proc. 28th Colloq., Peeters, H., ed., pp.51-56, Pe

A;Title: Human fibrinogen: sequence, sulfur bridges, glycosylation and some structural v

A;Reference number: A94433

A;Contents: carbohydrate binding

A;Accession: A94433

A;Molecule type: protein

A;Residues: 31-137,'QS',140-144,'QF',147-491 <HENS>

R;Watt, K.W.K.; Takagi, T.; Doolittle, R.F.

Biochemistry 18, 68-76, 1979

A;Title: Amino acid sequence of the beta chain of human fibrinogen.

A;Reference number: A90437; MUID:79124640; PMID:420779

A;Accession: A90437

A;Molecule type: protein

A;Residues: 31-144,'QF',147-231,'D',233-330,'E',332-491 <WAT>

R;Blomback, B.; Hessel, B.; Hogg, D.

Thromb. Res. 8, 639-658, 1976
A;Title: Disulfide bridges in NH-2-terminal part of human fibrinogen.
A;Reference number: A94309; PMID:76225080; PMID:936108
A;Contents: disulfide bonds
A;Accession: A94309
A;Molecule type: protein
A;Residues: 31-112, 'E', 114-137, 'QS', 140-144, 'QF', 147-148 <BLO>
R;Kunitake, S.T.; Carilli, C.T.; Lau, K.; Protter, A.A.; Naya-Vigne, J.; Kane, J.P.
Biochemistry 33, 1988-1993, 1994
A;Title: Identification of proteins associated with apolipoprotein A-I-containing lipoproteins
A;Reference number: A54223; PMID:94162201; PMID:8117655
A;Accession: G54223
A;Molecule type: protein
A;Residues: 164-174 <KUN>
R;Henschen, A.; Lottspeich, F.; Kehl, M.; Southan, C.
Ann. N. Y. Acad. Sci. 408, 28-43, 1983
A;Title: Covalent structure of fibrinogen.
A;Reference number: A90037; PMID:83254370; PMID:6575689
A;Contents: annotation; review, disulfide bonds
R;Gardlund, B.; Hessel, B.; Marguerie, G.; Murano, G.; Blombach, B.
Eur. J. Biochem. 77, 595-610, 1977
A;Title: Primary structure of human fibrinogen. Characterization of disulfide-containing
A;Reference number: A91249; PMID:77245999; PMID:891553
A;Contents: annotation; disulfide bonds
R;Doollittle, R.F.; Takagi, T.; Watt, K.; Bouma III, H.; Cottrell, B.A.; Cassman, K.G.; G
in Regulatory Proteolytic Enzymes and Their Inhibitors, Magnusson, S., Ottosen, M., Fold
A;Title: The structures of fibrinogen and fibrin.
A;Reference number: A94437
A;Contents: annotation; disulfide bonds
R;Doollittle, R.F.
Annu. Rev. Biochem. 53, 195-229, 1984
A;Title: Fibrinogen and fibrin.
A;Reference number: A90041; PMID:84305751; PMID:6383194
A;Contents: annotation; review, EM structure, polymerization, ligands
R;Chung, D.W.; Rixon, M.W.; Que, B.G.; Davie, E.W.
Ann. N. Y. Acad. Sci. 408, 449-456, 1983
A;Title: Cloning of fibrinogen genes and their cDNA.
A;Reference number: A90038; PMID:83254384; PMID:6575700
A;Contents: annotation
R;Kirschbaum, N.B.; Budzynski, A.Z.
J. Biol. Chem. 265, 13669-13676, 1990
A;Title: A unique proteolytic fragment of human fibrinogen containing the Aalpha COOH-ter
A;Reference number: A37117; PMID:90337977; PMID:2143188
A;Contents: annotation; hementin cleavage site
A;Note: hementin, a protease from Haemeteria ghilianii, the giant South American leech,
C;Comment: The conversion of fibrinogen to fibrin is triggered by thrombin, which cleave
ization sites responsible for the formation of the soft clot.
C;Comment: The soft clot is converted into the hard clot by factor XIIIa (fibrin-stabili
ger) and between alpha chains (weaker) of different monomers.
C;Comment: All fibrinogen chains are synthesized in the liver.
C;Genetics:
A;Gene: GDB:FGB
A;Cross-references: GDB:119130; OMIM:134830
A;Map position: 4q28-4q28
A;Introns: 38/3; 102/3; 164/1; 240/1; 278/1; 320/1; 415/2
C;Complex: The fibrinogen molecule is a hexamer containing two sets of alpha (see PIR:FG
ins are contained in the core. Two three-chain coiled coils emerge from this core and co
from the distal domain nodes.
C;Function:
A;Description: fibrinogen cleaved by thrombin yields monomers that are polymerized into
A;Pathway: blood coagulation
C;Superfamily: fibrinogen beta chain; fibrinogen beta/gamma homology; fibrinogen disulfid
C;Keywords: blood coagulation; coiled coil; glycoprotein; liver; plasma; pyroglyutamic ac
F;1-30/Domain: (or 4-30 or 15-30) signal sequence #status predicted <SIG>
F;31-491/Product: fibrinogen beta chain #status experimental <MAT>
F;31-44/Product: fibrinopeptide B #status experimental <APT>
F;45-491/Product: fibrin beta chain #status experimental <FGB>
F;45-47/Region: polymerization site
F;99-228/Domain: fibrinogen disulfide ring homology <PDR>
F;238-487/Domain: fibrinogen beta/gamma homology <FBG>
F;31/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status experimen
F;44-45/Cleavage site: Arg-Gly (thrombin) #status experimental

F;95/Disulfide bonds: interchain (to alpha-55) #status experimental
F;106/Disulfide bonds: interchain (to alpha-68) #status experimental
F;110/Disulfide bonds: interchain (to gamma-45) #status experimental
F;223/Disulfide bonds: interchain (to alpha-184) #status experimental
F;227/Disulfide bonds: interchain (to gamma-161) #status experimental
F;231-316,241-270,424-437/Disulfide bonds: #status experimental
F;394/Binding site: carbohydrate (Asn) (covalent) #status experimental

Query Match 99.8%; Score 2671; DB 1; Length 491;
Best Local Similarity 99.8%; Pred. No. 3.1e-170;
Matches 490; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MKRMVSWSFHKLKTKMKHLLLLCVFLYKSGQVNDNEEGFFSARGHRPLDKKREAPSRLR 60
DB 1 MKRMVSWSFHKLKTKMKHLLLLCVFLYKSGQVNDNEEGFFSARGHRPLDKKREAPSRLR 60
QY 61 PAPPISGGGYRPAKAAATQKVERKAPDAGGCHADPDGLVLCPTGCOLEALLQOE 120
DB 61 PAPPISGGGYRPAKAAATQKVERKAPDAGGCHADPDGLVLCPTGCOLEALLQOE 120
QY 121 RPIRNSVDELNNVEAVSQTSSSFQYMYLLKDLWKQKQKQVNDNENNVNYSSELEKHQ 180
DB 121 RPIRNSVDELNNVEAVSQTSSSFQYMYLLKDLWKQKQKQVNDNENNVNYSSELEKHQ 180
QY 181 LYIDTVNSNTATNLRLVRSILENLSKIQLKESDVSAQMEYCRTPCTVSCNIPVVSQKE 240
DB 181 LYIDTVNSNIPNLRLVRSILENLSKIQLKESDVSAQMEYCRTPCTVSCNIPVVSQKE 240
QY 241 CEIIRKGGTSEMYLIQPDSSVKPYRYCDMNTENGWTVIONRODGSVDGRKWDVPYK 300
DB 241 CEIIRKGGTSEMYLIQPDSSVKPYRYCDMNTENGWTVIONRODGSVDGRKWDVPYK 300
QY 301 QGFGNVATNTDCKNYCGLPGEYWLGNDKISQLTRMGPTLELLIEMEDWKDKYKAHYGGFT 360
DB 301 QGFGNVATNTDCKNYCGLPGEYWLGNDKISQLTRMGPTLELLIEMEDWKDKYKAHYGGFT 360
QY 361 VQNEANKYQISVNVKRGTAGNALMDGASQLMGENTMTTHNGMPFSTYDRDNDGLTSDP 420
DB 361 VQNEANKYQISVNVKRGTAGNALMDGASQLMGENTMTTHNGMPFSTYDRDNDGLTSDP 420
QY 421 RKQCKEDGGGWWYRNCHAAENGRYWGQYTDMAKHGTDGTVVMWVKGSWYSRMKM 480
DB 421 RKQCKEDGGGWWYRNCHAAENGRYWGQYTDMAKHGTDGTVVMWVKGSWYSRMKM 480
QY 481 SMKIRPFFPQQ 491
DB 481 SMKIRPFFPQQ 491

RESULT 2

FGBOB

fibrinogen beta chain - bovine
N;Contains: fibrinopeptide B
C;Species: Bos primigenius taurus (cattle)
C;Date: 29-Jul-1981 #sequence revision 29-Jul-1981 #text_change 13-Sep-1996
C;Accession: A03122; B03117; B37507; A37513; S02443
R;Blombach, B.; Doollittle, R.F.
Acta Chem. Scand. 17, 1816-1819, 1963
A;Title: The sequence of amino acids at the N-terminal end of bovine fibrinopeptide B.
A;Reference number: A03122
A;Accession: A03122
A;Molecule type: protein
A;Residues: 1-4 <BLO>
R;Sjoquist, J.; Blombach, B.; Wallen, P.
Ark. Kemi 16, 425-436, 1960
A;Title: Amino acid sequence of bovine fibrinopeptides.
A;Reference number: B03117
A;Accession: B03117
A;Molecule type: protein
A;Residues: 5-21 <SJO>
R;Martinelli, R.A.; Inglis, A.S.; Rubira, M.R.; Hageman, T.C.; Hurrell, J.G.R.; Leach, S.
Arch. Biochem. Biophys. 192, 27-32, 1979
A;Title: Amino acid sequences of portions of the alpha and beta chains of bovine fibrinope

Gene 75, 131-136, 1989


```

Query Match      23.6%; Score 630.5; DB 1; Length 437;
Best Local Similarity 33.6%; Pred. No. 1.9e-34;
Matches 142; Conservative 71; Mismatches 140; Indels 69; Gaps 15;

QY      95  CLHADPDLGVLCPTGC-----OLQEQALLQERPRPTNSVDELNNVQAV 137
Db      34  CCLIDERFGSYCTTTGGIADFLSTYQTKVDKDLQSLIEDILHQ---VENKTSYKVLQIKAI 90

QY      138  SQT----SSSSFOYMYLLKDLQKQKQKVDNNVNNVEYSSELEKHQ---LYIDETVNSNI 191
Db      91  QLTYPNDESSKPNMIDAATL--KSRKMLEE----IMKYEASILTHDSSRYLQEIYNSN- 143

QY      192  AYNLRVLRISILENLRSKIOKLESDDVSAQMEYCRTPCTVCSNIPVSGKCEBEIIRKGET 251
Db      144  ---NOKIV-----NLKEKVAQLAQ-----COEPCKDTVQIHDITGCKDQDILANKGAKQ 189

QY      252  SEMYLIQPDSSVKPRVYVYCDMTNENGWTVIQNRDGSVDGFRKWDPPYKQGFQNVQ-TWT 310
Db      190  SGLYPIKPLKANQQFLVYCEIGSGNGNGWTVFQKRDLGSDVDFKKQWITQYKEGFGHLSPTGT 249

QY      311  DGKNYCGLPGEYWLGNDKISQLTRMG--PTELLIEMEDWKDKVAHYGGFTVQNEANKY 368
Db      250  T-----EFWLGNKIKHLISTQSAIPYALRVELEDWNGRTSTADYAMFKVKGPEADKY 300

QY      369  QISVKNYKRTAGNALMDGASQLMGE--NRTWTIINGMFFSYTDRNDGWLITDPRKQCS 425
Db      301  RLTAYAFAGDAGDAFDGDFD--FGDDPSDKFTTSNGMQFSTWDNDNDKFL-----EGNCA 353

```


Query Match 22.8%; Score 609; DB 2; Length 444;
 Best Local Similarity 33.0%; Pred. No. 5.3e-33;
 Matches 134; Conservative 77; Mismatches 157; Indels 38; Gaps 12;

QY 95 CLHADPDLGLVPTGQLOEQALLQOERPIRNSVDELNNNVEAVSQTSSSFQMYLLKDL 154
 DB 32 CCILDERFGSCPTTCGIADELNNYQTSVDKOLRTLEGILYQENKTSBARELVKAIQIS 91

QY 155 WOKROKQVND-ENVNVEYSSELEKHQLYIDETVNSNIATNLRLVRSI-----LENLR 206
 DB 92 YNPDQSKNNIBSATYKNSKMMELMKY--ETLSTHETSTIRFLOEVYNSNSQKTVNLR 149

QY 207 SKIQKLESVSAQMEYCRTPCTVSCNIPVVSQKECEIIRKGETSEMYLIQDSSVKPY 266
 DB 150 DKVVQLEAN-----COEPCQDTXIHVDVTGEDCDVANKGAKESGLFIIRP-LKAKQF 201

QY 267 RYCDMNTENGWTVIQNRQDGSVDFGRKWDVPYKQFGVGNVATNDGKNCYGLPGEYWLGN 326
 DB 202 LVYCEIDGSGNGWTVFQKRLDGLSLDFKNNIQYKEGFHLSPTGTGNT-----EFWLGN 255

QY 327 DKISOLTRMG--PTELLIEMEDWKDKVKAHYGGFTVQNEANKYQISVKNY-RGTAGNAL 383
 DB 256 EXHILISTOSSIPYVIRIQLEMDNGRTSTADYASFKVTGENDKYRLTYAYFTGGDAGDA- 314

QY 384 MDGASQLMGE--NRTMTIHNGMFFSTYDRDNDGLTSDPRKQCKEDGGGWMYNNRCHAA 440
 DB 315 FDGYD--FGDDSSDKFPTSHNGMQFSTWSDNDKY-----DGNCAQVGLGWMWKNCHAG 367

QY 441 NPNRYWGGQYTWMAKHGTDGTVVMWKNKGSWYMRKWSMKIRP 486
 DB 368 HLNGVYVYQGYTSKTSPTNGYDNGIITWATWRKWSYMKKTKTKIIP 413

RESULT 9
 A32670
 fibrinogen gamma chain - African clawed frog
 C:Species: Xenopus laevis (African clawed frog)
 C>Date: 15-Jun-1990 #sequence_revision 15-Jun-1990 #text_change 09-Jul-2004
 C:Accession: A32670; I51416
 R:Pastori, R.L.; Moskaitis, J.E.; Smith Jr., L.H.; Schoenberg, D.R.
 Biochemistry 29, 2999-2605, 1990
 A:Title: Estrogen regulation of Xenopus laevis gamma-fibrinogen gene expression.
 A:Reference number: A32670; MUID:90241882; PMID:2334684
 A:Accession: A32670
 A:Molecule type: mRNA
 A:Status: preliminary
 A:Residues: 1-438 <PAS>
 A:Cross-references: UNIPROT:P17634; GB:J02894; NID:G214139; PIDN:AAA49709.1; PID:G214140
 R:Bhattacharya, A.; Shepard, A.R.; Moser, D.R.; Holland, L.J.
 Mol. Cell. Endocrinol. 72, 213-220, 1990
 A:Title: Isolation and characterization of cDNA clones for the gamma subunit of Xenopus
 A:Reference number: I51416; MUID:91146806; PMID:2289632
 A:Accession: I51416
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-58 <BHA>
 A:Cross-references: GB:M35548; NID:G214141; PIDN:AAA03247.1; PID:G214142
 C:Superfamily: fibrinogen gamma chain; fibrinogen beta/gamma homology
 C:Keywords: blood coagulation
 F:173-413/Domain: fibrinogen beta/gamma homology <FBG>

Query Match 22.7%; Score 607.5; DB 2; Length 438;
 Best Local Similarity 32.2%; Pred. No. 6.5e-33;
 Matches 137; Conservative 74; Mismatches 146; Indels 69; Gaps 14;

QY 90 PDAGGLHADPDLGLVPTGQLOEQALLQOERPIRNSVDELNNNVEAVS-QTSSSFQYM 148
 DB 27 PNTDNCILDRFGYBCTTCGTSDFLNRYQENVDTDLQYLENLLTQISNSTGTTIIVE 86

QY 149 YLL-----KDLWOKROKQVNDNENNVNVEYSSELEKHQLYIDETV 187
 DB 87 HLIDSGKKPATSPQTAIDPMTOKSKTCWKK-----LTDKKNY-----QYEENILYQEVY 137

QY 188 NSNIATNLRLVRSILLENLSKIQKLESVSAQMEYCRTPCTVSCNIPVVSQKECEIIRK 247
 DB 138 SSN-----QNKIFLLKOKIANLELQ-----CQPCRDVTQIOEFTGDKDCQEVANK 182

QY 248 GGETSEMYLIQDSSVKPYRVYCDMNTENGWTVIQNRQDGSVDFGRKWDVPYKQFGVNA 307
 DB 183 GARLSGLYYIKPLKAKQOFLVYCEIEPFSASMTVIQRLDGSVNFHKNVQVREGFGYLS 242

QY 308 TMTDGNKNCGLPGEYWLGNDKISQLTRMG--PTELLIEMEDWKDKVKAHYGGFTVQNEA 365
 DB 243 PND-----KTEFWLGNKEKIHLLSTQSTIPYVVRILELDSNQKSTADYSTFRLGSEK 294

QY 366 NKYQISVKNY-RGTAGNALMDGASQLMGE--NRTMTIHNGMFFSTYDRDNDGLTSDPR 421
 DB 295 DNYRFTYAYFIQGDAGDA--FDGFD--FGDDPSDKFYTSHNGMQFSTFDKNDKF-----D 346

QY 422 KQCKEDGGGWMYNNRCHAAANPNRGYVWGGQYTWMAKHGTDGTVVMWKNKGSWYMRK 480
 DB 347 GNCABQDGGWMWNNRCHAAHLNGKYYQGYTSEADSGPSGYDNGIITWATWRKWSYMKSV 406

QY 481 SMKIRP 486
 DB 407 TWKIMP 412

RESULT 10
 FGLMGS
 fibrinogen gamma chain precursor - sea lamprey
 C:Species: Petromyzon marinus (sea lamprey)
 C>Date: 17-Mar-1987 #sequence_revision 17-Mar-1987 #text_change 09-Jul-2004
 C:Accession: A03129
 R:Strong, D.D.; Moore, M.; Cottrell, B.A.; Bohonus, V.L.; Pontes, M.; Evans, B.; Riley, N
 Biochemistry 24, 92-101, 1985
 A:Title: Lamprey fibrinogen gamma chain: cloning, cDNA sequencing, and general characteri
 A:Reference number: A03129; MUID:85199776; PMID:2581603
 A:Accession: A03129
 A:Molecule type: mRNA
 A:Residues: 1-432 <STR>
 A:Cross-references: UNIPROT:P04115; GB:K03049; NID:G213193; PIDN:AAA49262.1; PID:G213194
 C:Superfamily: fibrinogen gamma chain; fibrinogen beta/gamma homology
 C:Keywords: blood coagulation; glycoprotein; liver; plasma; pyroglutamic acid
 F:1-24/Domain: signal sequence #status predicted <SIG>
 F:25-432/Product: signal sequence #status predicted <SIG>
 F:175-411/Domain: fibrinogen gamma chain #status experimental <MP>
 F:25/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status experiment
 F:227/Binding site: carbohydrate (Asn) (covalent) #status predicted
 F:421/Cross-link: isopeptide (Gln) (interchain to Lys-425 N6-amino) #status predicted
 F:425/Cross-link: isopeptide (Lys) (interchain to Gln-421) #status predicted

Query Match 22.4%; Score 598.5; DB 1; Length 432;
 Best Local Similarity 33.4%; Pred. No. 2.5e-32;
 Matches 144; Conservative 80; Mismatches 138; Indels 69; Gaps 17;

QY 84 KVERKAPDAGGLHADPDLGLVPTGQLOEQALLQOERPIRNSVDELNNNVEAVSQTSS 143
 DB 21 QVHAQVRDLKQCSN-DPEFGRYCPTTCGVADVLSKYAK-----GVDESSFISV----- 69

QY 144 SFQYMYLLKDLWOKROKQ--VKDNENNVNVEYSSELEKHQLYI-----DETNSNIATNL 196
 DB 70 -----LTQAAKHGIVEGNVIVNE-DVIRTRDEAQIKDSGQKTV-QKILEEVR 117

QY 197 VLRSILENLSKIQKLE-----SDVSAQMEYCRTPCTVSC-----NIPVVSQKE 240
 DB 118 ILQIGVSHDAQISEMWRVNOQFVTRLOQLVDIRQTCSRSQDTTANKISIPITGKD 177

QY 241 CEEIIRKGETSEMYLIQDSSVKPYRVYCDMNTENG--GWTVIQNRQDGSVDFGRKWDPY 299
 DB 178 CQQVNDGKSGSLYYIKPLKAKQOFLVYCEI--ENGNWTVIQHRHDSVNFTRDWSY 235

QY 300 KQFGNVATNTDGRKNCYGLPGEYWLGNDKISQLTRMGPTTELLIEMEDWKDKVKAHYGGF 359
 DB 236 REGGYLAPT-----LTTEFWLGNKEKIHLLTQQAQVRLRLDLTDWENTHRYADYGHF 287

JN0596
fibrinogen-related protein HFREP-1 precursor - human

RESULT 15
JN0596
fibrinogen

Search completed: November 22, 2004, 18:40:27
Job time : 43 secs

```

RX MEDLINE=21361164; PubMed=11468164;
RA Lounes K.C., Lefkowitz J.B., Henschen-Edman A.H., Coates A.I.,
Query Match 99.8%; Score 2671; DB 1; Length 491;
Best Local Similarity 99.8%; Pred. No. 2.3e-170;
Matches 490; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 MKRMVSHFHLKTMKHLHLLLLCVLVKSGVNDNEEGFFSARGHRPLDKKREAPSLR 60
Db 1 MKRMVSHFHLKTMKHLHLLLLCVLVKSGVNDNEEGFFSARGHRPLDKKREAPSLR 60
QY 61 PAPPIPGGGRARPAKAAATOKKVERKAPDAGGCLHADPDLGLVCTGCGLOEALLQOE 120
Db 61 PAPPIPGGGRARPAKAAATOKKVERKAPDAGGCLHADPDLGLVCTGCGLOEALLQOE 120
QY 121 RPIRNSVDELNNVEAVSOTSSSSFOYMYLLKDLWQKQKQVNDNENNVNVEYSSELEKHQ 180
Db 121 RPIRNSVDELNNVEAVSOTSSSSFOYMYLLKDLWQKQKQVNDNENNVNVEYSSELEKHQ 180
QY 181 LYIDETVNSIATNLRLVLRISILENLRKIQKLESQVSAQMEYCRTPCTVSCNIPVVSQKE 240
Db 181 LYIDETVNSIATNLRLVLRISILENLRKIQKLESQVSAQMEYCRTPCTVSCNIPVVSQKE 240
QY 241 CEIIRKGGTSEMYLIQPDSSVKPRVYCDMNTENGWTVIQNRQDGSVDFGRKWDPPYK 300
Db 241 CEIIRKGGTSEMYLIQPDSSVKPRVYCDMNTENGWTVIQNRQDGSVDFGRKWDPPYK 300
QY 301 QGFGNATNTDGNKCYGLPEYWLGNKDKISQLTRMGTELLIEMEDWKDGKVAHYGGFT 360
Db 301 QGFGNATNTDGNKCYGLPEYWLGNKDKISQLTRMGTELLIEMEDWKDGKVAHYGGFT 360
QY 361 VQNEANKYQISVKNYKRGTAGNALMDGASQVLMNGENRTWTIHNGMFFSTYDRDNDGWLTSQP 420
Db 361 VQNEANKYQISVKNYKRGTAGNALMDGASQVLMNGENRTWTIHNGMFFSTYDRDNDGWLTSQP 420
QY 421 RKQCKEDGGGWNYNRCHAANPNRYWGQYQYTWDMAKHGTDDGVVMMNKGWSYMRKM 480
Db 421 RKQCKEDGGGWNYNRCHAANPNRYWGQYQYTWDMAKHGTDDGVVMMNKGWSYMRKM 480
QY 481 SMKIRPFPQ 491
Db 481 SMKIRPFPQ 491
RESULT 2
Q8K0E8
ID Q8K0E8 PRELIMINARY; PRT; 481 AA.
AC Q8K0E8;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DE 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Fibrinogen, B beta polypeptide.
GN Name-Fgb;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=FVB/N; TISSUE=Liver;
RX MEDLINE=2238257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Murzyn D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
```

```

RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smallos D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RL and mouse cDNA sequences";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=FVB/N; TISSUE=Liver;
RA Strausberg R.;
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC031175; AAH31175.1; -.
DR HSSP; P02675; IFZE.
DR MGD; MGI:99501; Fgb.
DR InterPro; IPR002181; Fibrinogen C.
DR Pfam; PF00147; Fibrinogen_C; 1.
DR SMART; SM00186; FGB; 1.
DR PROSITE; PS00514; FIBRIN_AG_C_DOMAIN; 1.
SQ SEQUENCE 481 AA; 54752 MW; 9902830CF708A155 CRC64;
Query Match 84.4%; Score 2259; DB 2; Length 481;
Best Local Similarity 83.8%; Pred. No. 8.4e-143;
Matches 403; Conservative 42; Mismatches 32; Indels 4; Gaps 2;
QY 15 MKHL-LLLLLCLVPLVKSQVNDNEE---GFFSARGHRPLDKKREAPSLRPPPISSGG 70
Db 1 MRHLWLLLLCLVFSVQTQAADDYDPTDLSLDARGHRPVDKREKPEPSLRAPPPISSGG 60
QY 71 YRARPAAKATOKKVERKAPDAGGCLHADPDLGLVCTGCGLOEALLQOEPRPNSVDEL 130
Db 61 YRARPAAKATOKKVERKAPDAGGCLHADPDLGLVCTGCGLOEALLQOEPRPNSVDEL 120
QY 131 NNNEAVSOTSSSSFOYMYLLKDLWQKQKQVNDNENNVNVEYSSELEKHLYIDETVNSN 190
Db 121 NNNIQSVSDTSSTVTFQYLLTLKDKMKKQAQVKNENVINVEYSSILEDQELYIDETVND 180
QY 191 IATNLRLVLSILENLRKIQKLESQVSAQMEYCRTPCTVSCNIPVVSQKECEIIRKGG 250
Db 181 IPLNRLVLSILENLRKIQKLESQVSAQMEYCRTPCTVSCNIPVVSQKECEIIRKGG 240
QY 251 TSEMYLIQPDSSVKPRVYCDMNTENGWTVIQNRQDGSVDFGRKWDPPYKQGNVATNT 310
Db 241 TSEMYLIQPDTSIKPRVYCDMNTENGWTVIQNRQDGSVDFGRKWDPPYKQGNVATNE 300
QY 311 DGKNYCGLPGEYWLGNKDKISQLTRMGTELLIEMEDWKDGKVAHYGGFTVQNEANKYQI 370
Db 301 DAKKYCGLPGEYWLGNKDKISQLTRMGTELLIEMEDWKDGKVAHYGGFTVQNEANKYQV 360
QY 371 SVNKYRGTAGNALMDGASQVLMNGENRTWTIHNGMFFSTYDRDNDGWLTSQPRKQCKEDGG 430
Db 361 SVNKYRGTAGNALMDGASQVLMNGENRTWTIHNGMFFSTYDRDNDGWLTSQPRKQCKEDGG 420
QY 431 GWNYNRCHAANPNRYWGQYQYTWDMAKHGTDDGVVMMNKGWSYMRKMIRPFPQ 490
Db 421 GWNYNRCHAANPNRYWGGLYSWDMKHGTDDGVVMMNKGWSYMRKMIRPFPQ 480
QY 491 Q 491
Db 481 Q 481
RESULT 3
FIBB RAT
ID FIBB RAT STANDARD; PRT; 479 AA.
AC P14480;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Fibrinogen beta chain precursor [Contains: Fibrinopeptide B].
GN Name=Fgb;
OS Rattus norvegicus (Rat).
```


FT	TURN	77	79
FT	STRAND	81	84
FT	HELIX	86	112
FT	TURN	113	113
SQ	SEQUENCE	468 AA; 53340 MW; 2DED42F443AA4B37 CRC64;	

Query Match 79.8%; Score 2135; DB 1; Length 468;
Best Local Similarity 82.2%; Pred.No 1.6e-134;
Matches 379; Conservative 40; Mismatches 42; Indels 0; Gaps

Qy	31	QGVNDNEEGFFSARGHRPLDKKREEAPELRAPPPISGGYRARPAAAAATOKVKERKAPPA
Db	8	EQGDDRPKVGLGARHPDYDKKEEAPESLRVPPPISSGGYRARPATATVGKKVERKPE
Qy	91	DAGGCLHADPDVLGVCPTGCQLQEALLQOERPTRNSVDENLNNVAVSOTSSSSPOMYMI
Db	68	DADGCLHADPDVLGVCPTGCKLQDTLVROERPFRKSIEDLRNTVDVSVRTSSSTFOYTIT
Qy	151	LKDLWOKRKQVKDNNVNVNYSSELEKHOLIYDETVNSNIATNLRLVLSILENLRSLKI
Db	128	LKNMKGRQNQVDNNVNVNYSHLEKHOLIYDETVKNIPFKULVLSILENLRSLKI
Qy	211	KLESVSAQMEYCRTCTVSCNIPVVSGKECEBIIRKGGETSEMILIQDSVSKPYRVYCV
Db	188	KLESVDSTQMEYCRTCTVTTCNIPVVSGKECEKIIRNEGSETSEMILIQEDSSKPYRVY
Qy	271	DNKTENGHTVIQNRDGSDVDFGRKWDPYKQFGNVATNTDGKNYCGLPGEYWLGNDKI
Db	248	DMKTEKGWTVIQNRDGSDVDFGRKWDPYKQFGNIATNAEGKKYCGVPGEYWLGNDRIS
Qy	331	QLTRMGPTELLIEMEDWKGDVKAKHYGGFTVONEANKYOISVNYKRGTTAGNALMDGASOL
Db	308	QLTNMGPTKLLIEMEDWKGDKYALYEGFTVONEANKYQLSVKXKGTAGNALIEGASOL
Qy	391	MGENRTMTIHNGMFSTYDRDNDGLTSDPRKQCSKEDGGGWYNRCHAAPNPGRYYWG
Db	368	VGENRTMTIHNSMFFSTYDRDNDGWTTPDRKQCSKEDGGGWYNRCHAAPNPGRYYWG
Qy	451	QYTWDMAXHGTDGQVVMNMKGSWYSMRKMSKIRPFPPQQ 491
Db	428	AYTWDMAXHGTDGQVVMNMWQGSWYSMKMSKIRPYFPPEQ 468

RESULT 5
Q7TIME5
ID Q7TIME5 PRELIMINARY; PRT; 503 AA.
AC Q7TIME5
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DE 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE AC1-581 (Ab1-216) (Ab1-181).
OS Rattus norvegicus (Rat.).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RA Xu C.S., Li W.Q., Li Y.C., Chang C.F., Chai L.Q., Yuan J.Y.,
RA Yang K.J., Yan H.M., Zhao L.F., Ma H., Wang L., Wang S.F., Han H.P.,
RA Wang G.P., Shi J.B., Rahman S., Wang Q.N., Zhang J.B.,
RL Submitted (JUN-2003) to the EMBL/GenBank/DDBJ databases.
[2]
RP SEQUENCE FROM N.A.
RA Xu C.S., Li W.Q., Li Y.C., Wang S.F., Han H.P., Wang G.P., Chai L.Q.,
RA Yuan J.Y., Yang K.J., Yan H.M., Chang C.F., Zhao L.F., Ma H., Wang L.,
RA Shi J.B., Rahman S., Wang Q.N., Zhang J.B.;
RL Submitted (JUN-2003) to the EMBL/GenBank/DDBJ databases.
[3]
RP SEQUENCE FROM N.A.
RA Xu C.S., Li W.Q., Li Y.C., Yang K.J., Yan H.M., Chang C.F., Zhao L.F.,
RA Ma H., Wang L., Wang S.F., Han H.P., Wang G.P., Chai L.Q., Yuan J.Y.,
RA Shi J.B., Rahman S., Wang Q.N., Zhang J.B.;
RL Submitted (JUN-2003) to the EMBL/GenBank/DDBJ databases.
[3]

```

DR EMBL; AY321323; AAP86255.1; -.
DR EMBL; AY325153; AAP92554.1; -.
DR EMBL; AY325147; AAP92548.1; -.
DR InterPro; IPR002181; Fibrinogen_C.
DR Pfam; PF00147; Fibrinogen_C; 1.
DR SMART; SM00186; FBG; 1.
DR PROSITE; PS00514; FIBRIN AG C DOMAIN; 1.
SQ SEQUENCE 503 AA; 56612 MW; 02B6A1E05E7B5C12 CRC64;

Query Match 77.7%; Score 2080; DB 2; Length 503;
Best Local Similarity 80.7%; Pred. No. 8.5e-131;
Matches 369; Conservative 42; Mismatches 44; Indels 2; Gaps 1;

QY 34 NDNEGGFSARHRLDKKREAPSLRPAPPISGGYRARPAAKAAATQKKVERKAPDAG 93
DB 21 NSKRDLSTIARHRLPDRKEPPLRPAPPISGGYRARPAAKADGQKKVERKPPDAG 80

QY 94 GCLHADPDLGVLCPGTCQLEALLQERPIRNSVDLNNNVEAVSQTSSSFQYMYLLKD 153
DB 81 GCVHGDGMVLCPTGCELRLQTLNHERPIKNSIAELNSINSVSETSVTFQYLLTKD 140

QY 154 LWOKRQKQVKNENVNEYSSLEKHQYIDETVNSIATNLRLVLSILENLSKIQKLE 213
DB 141 MWKKQAQYKDNENVINEYSSILEQKGYIDETVNDNIPNLRLVLSILENLSKIQKLE 200

QY 214 SDVSAQMEYCRTPCTVSCNIPVVSQKECEIIRKGGTSEMYLIQPDSSVKPYRYCDMN 273
DB 201 SDLSAQTEYCHTPTVCNIPVVSQKECEIIRKGGTSEMYLIQPDTSKPYRYVCDMK 260

QY 274 TENGWTVIQRQDGSVDFGRKWDPKYQGFNVATNDGKNVCGLPGEYWLGNDKISQIT 333
DB 261 TENGWTVIQRQDGSVDFGRKWDPKYQGFNVATNDGKNVCGLPGEYWLGNDKISQIT 320

QY 334 RMGPTELLIEMEDWKGDKVKAHGGFTVONEANKYQISVKNYRGTAAGNMDGASOLMGE 393
DB 321 RIGPTELLIEMEDWKGDKVKAHGGFTVONEANKYQISVKNYRGTAAGNMDGASOLMGE 380

QY 394 NRTMTTHNGMFFSTYDRDNGMLTDPKQCKEDGGGWWYRCHAAFNPGRYWGQYT 453
DB 381 NRTMTTHNGMFFSTYDRDNGMVTDPKQCKEDGGGWWYRCHAAFNPGRYWGGLYS 440

QY 454 WDMAKHGTDGDDGVVWNNKGSWTSMRKMSMKIRPFPPQ 490
DB 441 WDMSKHGTDDGVVWNNKGSWTPRQHLQLQRP--PQ 475

RESULT 6
FIBB_CHICK STANDARD; PRT; 463 AA.
AC Q02020;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 01-OCT-2004 (Rel. 45, Last annotation update)
DE Fibrinogen beta chain precursor [Contains: Fibrinopeptide B]
DE (Fragment).
DE Name=FBG;
GN Gallus Gallus (Chicken).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OC NCBI_TaxID=9031;
OX [1]
RN SEQUENCE FROM N.A., AND SEQUENCE OF 1-13 AND 18-39.
RX MEDLINE=91182745; PubMed=2009266;
RA Weisbach L., Oddoux C., Procyk R., Grlenger G.;
RT "The beta chain of chicken fibrinogen contains an atypical thrombin
RL Biochemistry 30:3290-3294 (1991).
CC -!- FUNCTION: Fibrinogen has a double function: yielding monomers that
CC polymerize into fibrin and acting as a cofactor in platelet
CC aggregation.
CC -!- SUBUNIT: Hexamer containing 2 sets of 3 nonidentical chains
CC (alpha, beta and gamma), linked to each other by disulfide bonds.

```

```

-!- PTM: Conversion of fibrinogen to fibrin is triggered by thrombin,
which cleaves fibrinopeptides A and B from alpha and beta chains,
and thus exposes the N-terminal polymerization sites responsible
for the formation of the soft clot. The soft clot is converted
into the hard clot by factor XIIIa which catalyzes the epsilon-
(gamma-glutamyl)lysine cross-linking between gamma chains
(stronger) and between alpha chains (weaker) of different
monomers.
-!- SIMILARITY: Contains 1 fibrinogen C-terminal domain.
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
EMBL; M58514; AAA48770.1; -.
PIR; A38463; A38463.
PDB; 1BI3; X-ray; B/E=1-463.
PDB; 1MIJ; X-ray; B/E=1-463.
InterPro; IPR002181; Fibrinogen_C.
Pfam; PF00147; Fibrinogen_C; 1.
SMART; SM00186; FBG; 1.
PROSITE; PS00514; FIBRIN AG C DOMAIN; 1.
3D-structure; Blood coagulation; Direct protein sequencing;
Glycoprotein; Plasma; Platelet; Sulfation.
NON_TER 1 17 Fibrinopeptide B.
PEPTIDE <1 17 Fibrinogen beta chain.
CHAIN 18 463 Sulfotyrosine (By similarity).
MOD_RES 5 5 Cleavage (by thrombin; release
SITE 17 18 fibrinopeptide B).
DISULFID 69 69 Interchain (with alpha) (By similarity).
DISULFID 80 80 Interchain (with alpha) (By similarity).
DISULFID 84 84 Interchain (with gamma) (By similarity).
DISULFID 197 197 Interchain (with alpha) (By similarity).
DISULFID 201 201 Interchain (with gamma) (By similarity).
DISULFID 205 244 By similarity.
DISULFID 215 244 By similarity.
DISULFID 397 410 By similarity.
CARBOHYD 367 367 N-linked (GlcNAc...) (Potential).
SEQUENCE 463 AA; 52678 MW; 2044CD49BA79EC7B CRC64;
Query Match 65.2%; Score 1744; DB 1; Length 463;
Best Local Similarity 66.2%; Pred. No. 2.3e-108;
Matches 305; Conservative 70; Mismatches 76; Indels 10; Gaps 4;
QY 35 DNEEGF--FSARGHRLDKREAPSLRPAPPISGGYRARPAAKAAATQKKVERKA--- 89
DB 6 DNEEDSPQIDARAHRLDKRQAAATLRPVAPPISGTGYQRPFK---QDKAMKKGPII 62
QY 90 -PDAGGLHADPDLGVLCPGTCQLEALLQERPIRNSVDLNNNVEAVSQTSSSFQYM 148
DB 63 YPDAGGCKHPLDELGVLCPTGCELOTLKQEKTVKPELDELKDRVAKFSDTSTTWQYV 122
QY 149 YLLKDLWQKQKQVKNENVNEYSSLEKHQYIDETVNSIATNLRLVLSILENLSK 208
DB 123 NMIDNKLVTQKQKNDIILSEYNTMELHYNVYIKDNLNDNIPSSLRLVLRVIDSLHK 182
QY 209 IQKLESDVAQMEYCRTPCTVSCNIPVVSQKECEIIRKGGTSEMYLIQPDSSVKPYR 268
DB 183 IQKLENAIATQTDYCRSPCVASCNIPVVSQKECEIIRKGGTSEMYLIQPDFTFYRV 242
QY 269 YCDMTNTEGGWTVIQRQDGSVDFGRKWDPKYQGFNVATNDGKNVCGLPGEYWLGN 328
DB 243 YCDMTNTEGGWTVIQRQDGSVDFGRKWDPKYQGFNVATNDGKNVCGLPGEYWLGN 301
QY 329 ISQITRMGPTELLIEMEDWKGDKVKAHGGFTVONEANKYQISVKNYRGTAAGNMDGAS 388
DB 302 ISQITKIGTPKVLIEEMEDWNGDKVSAFYGGFTIHNEGKYLQSLVSNYKGNAGNLMGAS 361

```

Db 356 TGFTVQNEANKYQLSVSGYKGTAGNALMDGASQLKGENRWTWTHNGMFFSTFDRDNDGQ 415

Qy 417 TSDPRKCSKSDGGGWWYNRCHAANPNGRYYWGQYTWDMAKHGTTDDGWWVWVWVWGSWYS 476

Db 416 HSDPNKCSKSDGGGWWYNRCHAANPNGRYYWGQYTWDMAKHGTTDDGWWVWVWVWVWGSWYS 475

Qy 477 MKRMSKIRPPF 488

Db 476 MKRMSKIRPVF 487

RESULT 8

Q6NYEL PRELIMINARY; PRT; 485 AA.

ID Q6NYEL AC Q6NYEL

DT 05-JUL-2004 (TrEMBLrel. 27, Created)

DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)

DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)

DE Hypothetical protein fgb.

GN Name:fgb;

OS Brachydanio rerio (Zebrafish) (Danio rerio).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;

OC Cyprinidae; Danio.

OX NCBI_TaxID=7955;

RN [1]

RP SEQUENCE FROM N.A.

RC MEDLINE=22388257; PubMed=12477932;

RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D., Altschul S.F., Zeeberg B.H., Buetow K.H., Schaefer C.F., Bhat N.K., Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F., Diatchenko L., Marudina K., Farmer A.A., Rubin G.M., Hong L., Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E., Brownstein M.J., Usdin T.B., Tothiyuki S., Carninci P., Prange C., Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J., Rosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H., Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W., Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Fahey J., Helton E., Kettaman M., Madan A., Rodrigues S., Sanchez A., Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G., Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butlerfield Y.S., Krzywinski M.I., Skalska U., Smalls D.E., Schnerch A., Schein J.E., Jones S.J., Marra M.A.;

RT "Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences.";

RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).

RN [2]

RP SEQUENCE FROM N.A.

RC TISSUE=Kidney;

RA Strausberg R.;

RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.

DR EMBL; BC066629; AAH66629.1; -

DR InterPro; IPR002181; Fibrinogen_C.

DR Pfam; PF00147; Fibrinogen_C; 1.

DR SMART; SM00186; FBG; 1.

DR Hypothetical protein.

SW SEQUENCE 485 AA; 54407 MW; A3622504448B0FF3 CRC64;

Query Match 54.7%; Score 1464; DB 2; Length 485;

Best Local Similarity 55.0%; Pred. No. 1.3e-89;

Matches 270; Conservative 83; Mismatches 118; Indels 20; Gaps 8;

Qy 15 MKHLILLLLC-VFLVKSQGVNDNEGFSA-----RGRPLDKGREAPSLRAP---P 64

Db 1 MKLVILLLLCLCAVGALLAQDDYDDYGEKKEAKEVVDPRGRHPVSRGRE---TSPGPVSP 57

Qy 65 PISGG-GYBARPAKA---AATQKVKERKAPDAGGCLHADPDGLVPTGQLQELALQOE 120

Db 58 PISGGTRVGRPTAAPVGVKAVKEKESQ---PESGCGNHMGKMGVLCPTGELKKALIKOE 115

KW	Hypothetical protein.
SQ	SEQUENCE 485 AA; A362250448B0FF3 CRC64;
	Query Match 54.7%; Score 1464; DB 2; Length 485;
	Best Local Similarity 55.0%; Pred. No. 1.3e-89;
	Matches 270; Conservative 83; Mismatches 118; Indels 20; Gaps 8
QY	15 MKHLLILLLC-VFLIVKSGQVNDNEEGFSA-----RGRPLDKKRKEEAPSLRPAP--P 64
DB	1 MKLVLLILCLCAVGALAAODDDYDGEKKAEKVDPGRHPVSRGRE---TVSPGPVPSP 57
QY	65 PISGG-CYRARPAKA--AATQKKVKRAPADAGCLHADPDIGVLCPGTCQLQEALLQQE 120
DB	58 PISGGTYRRGRPTAAPGVKAVQEKBEQ--PESGGCNHMSEKMGVLCPGTCECLKALKIQE 115
QY	121 RPIENSVDLENNNVEASVSOTSSSFQVMYLKLDLWKQPKQVKQKDENVNVYESSLEKHQ 180
DB	116 RNVPKPTVEQLKRAVDUUTQSTNSIHGYVLDMTAEVAQRQKYGSEGNGLVDQYDTSLETCH 175
QY	181 LYIDETVNSNIATNLRLVRSLILENLRSKIQLSDVSAQMERYCTPTVSCNIPVVSGKE 240
DB	176 AYIKDTVDVTFPONIKVLQGVLDKIREKIQRLEKAITTQAKCQACPCKVCTPIPVVSGKE 235
QY	241 CEEIIRRGGETSEMVLTOPDSSVXPYRVYCDMNTENGWTVIQNQDGSDVDFGRKWDPYK 300
DB	236 CEDIIIRKGGEDSQMYIIRPDIPLGTPIKYFCDOQTKNCGWLLIQNRMDGSVDFGRRWDDR 295
QY	301 QGFNGVATNTDGKNYCGLPGEYVTLGNDRIQSILTRMGPTELLIMEDWDGDKVKAHYGGFT 360
DB	296 RGFGNIAFDV-KGHCQTPGEYVWLGNDRISQLSKMGATELIVEMEDWSGKVYAQYEQFS 354
QY	361 VQNFANKYQISVNKVRGTAGNALMDGASQLMGENTMTIHNGMPFSTYDRDNDGWLTSOP 420
DB	355 MQGEASNYYILGVRYSGTAGNTPLEGATELFGENRTMTIHNGMWPFSTYDRDNDKWIPGDP 414
QY	421 RKQCSKEDGGGWYNRCHAANPNGRYYWGGOYTDMAXHGTDGDDGVVNMWKGSWYSMRK 480
DB	415 SKQCSKEDGGGWYNRCHSNPNGRYYWGGAITYKMXHGITDDGIWMWNMKGWSYLSKTI 474
QY	481 SMKIRPPFPQ 491
DB	475 SMKIRPYFKQK 485

296	RGFGNIAFDV-GKGHCQTPGGYIWLGNDRISQLSKMGATELIVEMEDWSGSKVYAQYEQFS	420
361	VONFANKYQISVNNKYRGTAGNALMDGASQLMGENTMTIHNGMFSTSYDRDNDGWLTSDP	420
355	MQGASNTILGVGRYSFGTAGNTFEGATLFGENTMTIHNGMFSTSYDRDNDKWIPGDP	414
421	RKQSKEDGGGWNNRCHAAHPNGRYGGQYTDMAKHGTDDGVVNNWNGKGSWYNNRM	480
415	SKQCSKEDGGGWNNRCHSCNPNGRYGGGATYKMAKHGTDDGIVNNWNGKGSWYSLKTI	474
481	SMKIRPPPPQQ 491	
475	SMKIRPYFQK 485	
RESULT 10		
Q7SYN8	PRELIMINARY; PRT; 321 AA.	
ID	Q7SYN8	
AC	Q7SYN8;	
DT	01-OCT-2003 (TREMBLrel. 25, Created)	
DT	01-OCT-2003 (TREMBLrel. 25, Last sequence update)	
DT	01-MAR-2004 (TREMBLrel. 26, Last annotation update)	
DE	LOC397825 protein (Fragment).	
GN	Name=LOC397825;	
OS	Xenopus laevis (African clawed frog).	
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
OC	Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;	
OC	Xenopodinae; Xenopus.	
OX	NCBI_TaxID=8935;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RC	TISSUE=Whole;	
EX	MEDLINE=23288257; PubMed=12477932;	
RA	Klausner R.L., Feingold E.A., Grouse L.H., Derge J.G.,	
RA	Stausberg R.D., Collins F.S., Wagner L., Shennen C.M., Schuler G.D.,	
RA	Aeschul S.F., Jordan B., Buetow K.H., Schaefer C.F., Bhat N.K.,	
RA	Hopkins R.F., Zeeberg H., Moore T., Max S.I., Wang J., Hsieh F.,	
RA	Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,	
RA	Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,	
RA	Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,	
RA	Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,	
RA	Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,	
RA	Richards S., Worley K.C., Hale S.J., Garcia A.M., Gay L.J., Hulyk S.W.,	
RA	Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,	

RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whitting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
 RA Krzywinski M.I., Skalska U., Smallus D.E., Schnerch A., Schein J.E.,
 RA Jones S.J., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Whole;
 RX MEDLINE=22341132; PubMed=12454917;
 RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
 RA Richardson P.;
 RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus
 RT initiative.";
 RL Dev. Dyn. 225:384-391(2002).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Whole;
 RA Klein S., Strausberg R.;
 RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC054326; AAH54326.1; -;
 DR InterPro; IPR002181; Fibrinogen_C.
 DR Pfam; PF00147; Fibrinogen_C; 1.
 DR SMART; SM00186; FBG; 1.
 DR PROSITE; PS00514; FIBRIN_AG_C_DOMAIN; 1.
 FT NON TER 1
 SQ SEQUENCE 321 AA; 36519 MW; E0C41FCC3EB40DEF CRC64;
 Query Match 51.7%; Score 1384.5; DB 2; Length 321;
 Best Local Similarity 75.4%; Pred. No. 1.6e-84;
 Matches 242; Conservative 33; Mismatches 45; Indels 1; Gaps 1;
 QY 168 VNEYSELEKQLYIDETVNSNIATNLRLVLSILENLRSKIQLSDVSAQMEYKRTPC 227
 DB 1 VVNEYNLEEEQYTFIKDNIDTKIPSNIRILRQVLENLSKIKLETAIATQVENCSPC 60
 QY 228 TVSCNIPVVGKCECEIIRKGGTSEMYLIQDSSVKKPYRVYCDMNTENGWTVIQNRQD 287
 DB 61 VTTCPVVGKCECEIIRKGGTSEMYLIQDSSVKKPYRVYCDMATHDGGWTVIQNRQD 120
 QY 288 GSVDFGRKWDPKYKQGFNGVATNTDGNKYCGLPGEYWLGNDKISQLTRMGPTTELLIEMEDW 347
 DB 121 GSVGFGRTWDSYSGFGNTAAN-GKGKICDMGFEFNLGNEKISQTLNLCATLAFEMEDW 179
 QY 348 KGDVKVKAHYGGFTVQNEANKYQISVKNKYRGTAGNALMDGASQLMGENRTMTIHNGMFFST 407
 DB 180 DGAKVTAQYTGFTVQNEANKYQISVKNKYRGTAGNALMDGASQLMGENRTMTIHNGMFFST 239
 QY 408 YDRDNDGMLTSDPRKQCKEDGGWYNNRCHAANPNRGYVGGQYTWDMAKHGTDDGVVW 467
 DB 240 FDRDNDGWHQSDPNKQCKEDGGWYNNRCHAANPNRGYVGGYTWDMAKHGTDDGVVW 299
 QY 468 MNWKGWSYMRKXSMKIRPF 488
 DB 300 MNWKGWSYMRKXSMKIRPF 320
 RESULT 11
 Q91ZP1 PRELIMINARY; PRT; 236 AA.
 AC Q91ZP1;
 DT 01-DEC-2001 (TREMELrel. 19, Created)
 DT 01-DEC-2001 (TREMELrel. 19, Last sequence update)
 DE 01-WAR-2004 (TREMELrel. 26, Last annotation update)
 GN Fibrinogen B-beta-chain (Fragment).
 OS Name=FBG;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NCBI_TaxID=10090;

RN SEQUENCE FROM N.A.
 RP Murakawa M., Freeman M.W.;
 RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF413205; AAL02225.1; -;
 DR MGD; MGI:99501; Fgb.
 DR InterPro; IPR002181; Fibrinogen_C.
 DR Pfam; PF00147; Fibrinogen_C; 1.
 DR SMART; SM00186; FBG; 1.
 DR PROSITE; PS00514; FIBRIN_AG_C_DOMAIN; 1.
 FT NON TER 1
 SQ SEQUENCE 236 AA; 27054 MW; A844F0E40CCB8B01 CRC64;
 Query Match 47.8%; Score 1279; DB 2; Length 236;
 Best Local Similarity 92.4%; Pred. No. 1.2e-77;
 Matches 218; Conservative 13; Mismatches 5; Indels 0; Gaps 0;
 QY 244 IIRKGGTSEMYLIQDSSVKKPYRVYCDMNTENGWTVIQNRQDGSVDFGRKWDPKYKQGF 303
 DB 1 IIRKGGTSEMYLIQDSSVKKPYRVYCDMNTENGWTVIQNRQDGSVDFGRKWDPKYKQGF 60
 QY 304 GNVAITNTDGNKYCGLPGEYWLGNDKISQLTRMGPTTELLIEMEDWKGDKVKAHYGGFTVQN 363
 DB 61 GNIATNEDAKYCGLPGEYWLGNDKISQLTRMGPTTELLIEMEDWKGDKVKAHYGGFTVQN 120
 QY 364 EANKQISVKNKYRGTAGNALMDGASQLMGENRTMTIHNGMFFSTYDRDNDGMLTSDPRKQ 423
 DB 121 EASKYQVSVNKKYGTAGNALMDGASQLMGENRTMTIHNGMFFSTYDRDNDGMLTSDPRKQ 180
 QY 424 CSKEDGGGWWYNNRCHAANPNRGYVGGQYTWDMAKHGTDDGVVWNNWKGWSYMRK 479
 DB 181 CSKEDGGGWWYNNRCHAANPNRGYVGGYTWDMKXHGTDGTVVNNWKGWSYMRK 236
 RESULT 12
 FIBB PETWA
 ID FIBB PETWA STANDARD; PRT; 477 AA.
 AC P02678;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 01-OCT-1989 (Rel. 12, Last sequence update)
 DT 01-OCT-2004 (Rel. 45, Last annotation update)
 DE Fibrinogen beta chain [Contains: Fibrinopeptide B] (Fragments).
 OS Petromyzon marinus (Sea lamprey).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Hyperoartia;
 OC Petromyzontiformes; Petromyzontidae; Petromyzon.
 NCBI_TaxID=7757;
 RN [1]
 RP SEQUENCE OF 1-36.
 RX MEDLINE=7065679; PubMed=999898;
 RA Cottrell B.A., Doolittle R.F.;
 RT "Amino acid sequences of lamprey fibrinopeptides A and B and
 RT characterizations of the junctions split by lamprey and mammalian
 RL thrombins.";
 RL Biochim. Biophys. Acta 453:426-438(1976).
 RN [2]
 RP SEQUENCE OF 37-477 FROM N.A.
 RX MEDLINE=87076582; PubMed=3790537;
 RA Bohonus V.L., Doolittle R.F., Pontes M., Strong D.D.;
 RT "Complementary DNA sequence of lamprey fibrinogen beta chain.";
 RL Biochemistry 25:6512-6516(1986).
 CC -!- FUNCTION: Fibrinogen has a double function: yielding monomers that
 CC polymerize into fibrin and acting as a cofactor in platelet
 CC aggregation.
 CC -!- SUBUNIT: Hexamer containing 2 sets of 3 nonidentical chains
 CC (alpha, beta and gamma), linked to each other by disulfide bonds.
 CC -!- PTM: Conversion of fibrinogen to fibrin is triggered by thrombin,
 CC which cleaves fibrinopeptides A and B from alpha and beta chains,
 CC and thus exposes the N-terminal polymerization sites responsible
 CC for the formation of the soft clot. The soft clot is converted
 CC into the hard clot by factor XIIIa which catalyzes the epsilon-
 CC (gamma-glutamyl)lysine cross-linking between gamma chains
 CC (stronger) and between alpha chains (weaker) of different

monomers.
 1- SIMILARITY: Contains 1 fibrinogen C-terminal domain.

 This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

 EMBL; M14773; AAA9261.1; -
 DR PIR; A25052; A25052.
 DR PDB; 1LWU; X-ray; B/E/H/K=155-477.
 DR PDB; 1N73; X-ray; B/E=155-477.
 DR InterPro; IPR002181; Fibrinogen_C.
 DR SMART; SM00186; FBG; 1.
 DR PROSITE; PS00514; FIBRIN AG C DOMAIN; 1.
 KW 3D-structure; Blood coagulation; Direct protein sequencing;
 KW Glycoprotein; Plasma; Sulfation.
 FT PEPTIDE 1 36 Fibrinopeptide B.
 FT NON_CONS 36 37
 FT CHAIN <37 477 Fibrinogen beta chain.
 FT MOD_RES 13 13 Sulfotyrosine.
 FT CARBOHYD 27 27 N-linked (GlcNAc...)
 FT DISULFID 84 84 Interchain (with alpha chain) (By similarity)
 FT DISULFID 95 95 Interchain (with alpha chain) (By similarity)
 FT DISULFID 99 99 Interchain (with gamma chain) (By similarity)
 FT DISULFID 212 212 Interchain (with alpha chain) (By similarity)
 FT DISULFID 216 216 Interchain (with gamma chain) (By similarity)
 FT DISULFID 220 304 By similarity.
 FT DISULFID 230 259 By similarity.
 FT DISULFID 412 425 By similarity.
 FT HELIX 162 168 By similarity.
 FT TURN 169 169
 FT HELIX 170 177
 FT TURN 178 180
 FT HELIX 181 188
 FT TURN 189 189
 FT HELIX 190 210
 FT TURN 211 212
 FT STRAND 215 215
 FT STRAND 217 218
 FT STRAND 222 223
 FT STRAND 227 227
 FT HELIX 230 235
 FT TURN 236 237
 FT STRAND 242 246
 FT TURN 250 251
 FT STRAND 255 260
 FT STRAND 263 265
 FT STRAND 268 274
 FT HELIX 285 290
 FT STRAND 291 293
 FT STRAND 296 297
 FT STRAND 306 307
 FT STRAND 310 312
 FT HELIX 314 323
 FT STRAND 326 333
 FT TURN 335 336
 FT STRAND 339 347
 FT STRAND 349 349
 FT STRAND 352 354
 FT HELIX 355 356
 FT TURN 358 358
 FT STRAND 361 361
 FT STRAND 370 373
 FT HELIX 370 373
 FT TURN 376 377

FT TURN 380 380
 FT HELIX 381 384
 FT TURN 385 386
 FT TURN 389 390
 FT STRAND 394 395
 FT TURN 396 397
 FT STRAND 398 398
 FT TURN 405 406
 FT TURN 408 409
 FT TURN 412 417
 FT STRAND 420 420
 FT STRAND 428 429
 FT TURN 430 431
 FT TURN 436 437
 FT STRAND 439 439
 FT TURN 442 444
 FT TURN 446 447
 FT STRAND 454 455
 FT HELIX 456 459
 FT TURN 461 462
 FT STRAND 463 463
 FT STRAND 467 473
 SQ SEQUENCE 477 AA; 54270 MW; B8A95E7E32D09D18 CRC64;
 Query Match 45.3%; Score 1211; DB 1; Length 477;
 Best Local Similarity 48.7%; Pred. No. 1.1e-72;
 Matches 219; Conservative 74; Mismatches 143; Indels 14; Gaps 4;
 QY 49 LDKKEEAPSLRPAPP-----ISGGVRRAPAKAAATQKKVERKAPDAGGCLHAPDGLG 103
 DB 33 LDVRLPESGTVRRPPPLRHRRLAPGAVMSRPPASPRPQEAQKAIREDGGCMLPESDLG 92
 QY 104 VLCPTGCOLQEAALQOERPINSVDLANN-----VEAVSQTSSTSSSFQYWLKLDWQKQ 159
 DB 93 VLCPTGCELRELLKQRPVRYKISMLKQNLTYFINSFDRMASDS-----NTLKQNVQTLR 148
 QY 160 KQVKNNVNVVSESELEKHQIYIDETVNSNTATNLRLVLRSTLEMLRSKIOKLESVSAQ 219
 DB 149 RLNSRSSSTHVNAAQEIENRYKEVKIRIESTVAGSLRSKMSVLEHLRAKQQRMEEAIKTQ 208
 QY 220 MEYCRTPCTVSCNIPVSGKECEIIRKGGTSEMYLIQPDSSVKPYRYVCDMNTENGW 279
 DB 209 KELCSAPCTVNCRVFVSGMHCEDIYRNGRTSEAYIOPDLFSEPKYKFCDMESHGGW 268
 QY 280 TVIQNRQDGSVDGRKWDYKQGFQGNVATNTDGKNYCGLPGEYWLGNDKISQITRMGPTE 339
 DB 269 TVVQNRVDGSSNFARDWNTYKAEGFNIAFG-NGKSIICNIPGEYWLGTQVHQLTKQHTQ 327
 QY 340 LLIEMEDWKGDKVAHYGOFTVQNEANKYQISVKNYRGTAGNALMDGASQALMGENTMTI 399
 DB 328 VLFDMSDWEGSSVYAYASFRPENEAGYRLWVEDYSGNAGNALLLEGATQLMGDNKMTI 387
 QY 400 HNGMFFSTYDRDNDGWLTSDFPRKQCKEDGGGWYNNRCHAAHPNGRYWGGQYTWDMAXH 459
 DB 388 HNGMQFSTFDRDNDNWNPGDPTKHCSDREDAGGWYNNRCHAAHPNGRYWGGQYTWDMAXH 447
 QY 460 GTDDGVVWNNKGSWYSMKRMSKIRPFP 489
 DB 448 GTDDGVVWNNKGSWYSMRQAMKLRPKWP 477
 RESULT 13
 Q7ZVG7
 ID Q7ZVG7 PRELIMINARY; PRT; 431 AA.
 AC Q7ZVG7
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Zgc:56023.
 GN Name=zgc:56023;
 OS Brachydanio rerio (Zebrafish) (Danio rerio).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;

	Cyprinidae; Danio.	
OC	NGBI_TaxID=7955;	
OX	[1]	
RN	SEQUENCE FROM N.A.	
RP	STRAIN=AB; TISSUE=Whole body;	
RC	MEDLINE=22388257; PubMed=12477932;	
RX	Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,	
RA	Klausner R.D., Collins F.S., Wagner I., Shermen C.M., Schuler G.D.,	
RA	Altschul S.F., Zeeberg B., Moore T.K., Max S.I., Wang J., Hsieh P.,	
RA	Hopkins R.F., Jordan H., Butte T.X., Kachur A.J., Rubin G.M., Hong L.,	
RA	Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,	
RA	Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,	
RA	Brownstein M.J., Usdin T.B., Toshikiuki S., Carninci P., Prange C.,	
RA	Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,	
RA	Besak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,	
RA	Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,	
RA	Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Sanchez A.,	
RA	Fahay J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,	
RA	Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,	
RA	Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,	
RA	Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,	
RA	Krzywinski M.I., Skalska U., Smallos D.E., Schnercher A., Schein J.E.,	
RA	Jones S.J., Marra M.A.;	
RT	"Generation and initial analysis of more than 15,000 full-length human	
RL	and mouse cDNA sequences.";	
RL	Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).	
RN	[2]	
RP	SEQUENCE FROM N.A.	
RC	STRAIN=AB; TISSUE=Whole body;	
RL	Strausberg R.	
RL	Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.	
DR	EMBL; BC045868; AAH45868.1; -.	
DR	HSSP; P02679; 2PiB.	
DR	InterPro; IPR002181; Fibrinogen_C.	
DR	Pfam; PF00147; Fibrinogen_C_1.	
DR	SMART; SMG0186; FBG; 1.	
DR	PROSITE; PS00514; FIBRIN_AG_C_DOMAIN; 1.	
SQ	SEQUENCE 431 AA; 48832 MW; 101B2EA7E33DFF71 CRC64;	
	Query Match 24.6%; Score 657; DB 2; Length 431;	
	Best Local Similarity 35.2%; Pred. No. 1,2e-35;	
	Matches 141; Conservative 85; Mismatches 127; Indels 48; Gaps 13;	
QY	103 GVLCPTGCQLQEALLQQERPIRNVSDELNNVFAVSTSSSFQYYLLKDLWQRKQKQ- 161	
Ddb	39 GTVCPTCGVADLYQRYKPDMDKKLDMEODLEEIANLTRAQDKVVYLKSEAQAQKS 98	
QY	162 ----VKDNENNVN-----EYSSELEKHQLYIDEVTNSNIATNLRVLSILENLSKIQL 212	
Ddb	99 PDTYIKSKSNMLDDILRFEXSLIAQEQIY-----QIQSVLOANEKKITDL 144	
QY	213 ESDEVSAQM-EYC RTPCTVCSNPVPWSKECEEIIIRKGGETSEMILI QPDS SVKPYRVYCD 271	
Ddb	145 K-QMSQLDQMCKEPCKDTVEIQTVTKDCQDIANKGGKVGSLGYVYPAPAEPAFLVYCE 203	
QY	272 MATTEGGWTVIQNROGSVDGFGRKNDPYKQGFGNAIVTDTGNKYCGLPGBYVLGN DKISQ 331	
Ddb	204 IDSFRGWTVLQRRRDSDVFSGKNWTQYKEGFYL--SPDDR-----EFWLGNBKHL 255	
QY	332 LTRMG--PTELTIEMDWKDGVKAHYGGFTVQNEANKYOISVNKY-RGTAGNALWDGAS 398	
Ddb	256 LSVQSQSVPVVLRLEMVDWEGNKKYADYATFKLGPEVDVAFLTAYTFGGDAGA-FDGYD 314	
QY	389 QLNGE---NRMTIHGMFFSTVDRNDGWLTSDDRKCQSKEDGGGWYNRCHAANPNGR 445	
Ddb	315 --FGDDPSDFKFTSHNGMQFSIVDRNDKF----QGHCAQQDGSGWNVRCHAHNGK 367	
QY	446 YYWGCGYTWDMAKHGTDGDGVVNWNGSWYSMKRMKMKNRP 486	
Ddb	368 YTYGGKYTKDAESGYDNGIILWATHSRWYSLKETTAKIP 408	

RESULT 14

```

Q93568          PRELIMINARY;             PRT;   435 AA.
ID O93568;
AC O93568;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Fibrinogen gamma chain precursor.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=White Leghorn; TISSUE=Liver;
RA Cao Y., Weisbach L., Fu Y., Grieninger G.;
RL Submitted (AUG-1998), to the EMBL/GenBank/DBJ databases.
DR EMBL; AF087432; AAC36476.1; -.
DR PDB; 1EI3; X-ray; C/F=27-435.
DR PDB; 1MLU; X-ray; C/F=27-435.
GO; GO:0003676; F:nucleic acid binding; IEA.
DR InterPro; IPR002181; Fibrinogen_C.
DR InterPro; IPR008994; Nucleic_acid_OB.
DR Pfam; PF00147; Fibrinogen_C; 1.
DR SMART; SM00186; FBG; 1.
DR PROSITE; PS00514; FIBRIN_AG_C_DOMAIN; 1.
KW Signal.
FT CHAIN
SQ SEQUENCE      1    26      Potential.
                        27  435      fibrinogen gamma chain.
                               435 AA; 49642 MW; 91D2BC7FE8E0766 CRC64;

Query Match           24.5%; Score 656.5; DB 2; Length 435;
Best Local Similarity 33.2%; Pred. No. 1.3e-35;
Matches 139; Conservative 76; Mismatches 141; Indels 63; Gaps 11;

Qy 95 CLHADPDLGVLCPTGCQLQ-----EALLQBRPIRNSVDELNNVEAV 137
Db | | | | | | | | | | | | | | | | | | | | | | : : : :
34 CCILDERFGSYCPTCGIADFFNKYRLTTDGELLEIEGLLQQATNSTGSLEYLIQHIKTI 93

Qy 138 ----SQTSSSFQYMVLKLDLWOKROKVQKDENVNVEYSSELEKHQLYIDETVNSNIAT 193
Db | : : : : | : : : : | : : : : | : : : : | : : : :
94 YPEKQTLPOS-----IEQLTKSKKIIE-----IIRVENTTILAHTNTQQLTDMHNM 143

Qy 194 NLRVLSILENRASKIQLESVDVSAQMEYCRTPTCVSCNPVPVSKECEELIRKGGETSE 253
Db | : : : : | | | | | | | | | | | | | | | | | : : : :
144 SNKITQ-----LKQKTAQLS-----HCQEFCKDTAELOETIGRDCQDIANKGARKSG 191

Qy 254 MYLIQPDSSVKPVRYCDMNTENGMTVIQNROGDGVDFGRKWDPPYKQGFGNVATTNDGK 313
Db | : : : : | : : : : | : : : : | : : : : | : : : :
192 LYFIKPQAKQSPLVYCEIDTYGWTGLQRLLDGEDFRFNWVQYKEGFGLSPDD--- 248

Qy 314 NYCGLPGEYWLGNDKISQLTRMG--PTELLIEWEDWKGVKAHYGGFVTQNEANKYQIS 371
Db | : : : : | : : : : | : : : : | : : : : | : : : :
249 -----TTEFWLGNEKLHLITOSTLPVALRIELEDSWGKKGTADYAVFKVGTEEDKYRLT 303

Qy 372 VNKYRGTAGNALMDGASQLMGE---NRTMTIHNGMPFFSTYDRDNCGWLTS DPRKCQSKED 428
Db | : : : : | : : : : | : : : : | : : : : | : : : :
304 YAYFIGERGDAFDGFN--FGDDPSDKSYTINGMRFSFTFDNDNF-----EGNCAREQD 356

Qy 429 GGGWYNRCHAAHPNGRYVYGQYTWMMAHGTDGVMWNMKGSWYSRMKMSMKIRPF 487
Db | : : : : | : : : : | : : : : | : : : : | : : : :
357 GSGWMMNRCHAGHLNGPYIGVSRDTCGNSYDNGIIWATWRDRWSYMKKTMTKIIPF 415


RESULT 15
Q8VCW7          PRELIMINARY;             PRT;   436 AA.
ID Q8VCW7
AC Q8VCW7; Q8WUR3;
DT 01-WAR-2002 (TrEMBLrel. 20, Created)
DT 01-WAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Fibrinogen, gamma polypeptide.
BN Name=Fbg;

```

Query Match	23.68;	Score	631.5;	DB	2;	Length	436;		
Best Local Similarity	33.5%;	Pred. No.	6e-34;						
Matches	140;	Conservative	79;	Mismatches	138;	Indels	61;	Gaps	15;
QY	95	CLHADPDLGVLCPTGCQLQEAALLQOERPIRNSVDELANNV-EAVSQTSSS-----SQYQM	148						
DB	33	CCILDERFGSPCTTCGADFLSSQYQTDVNDLTLSDILFRAENRTEAKELIKALQVY	92						
QY	149	Y-----LLKLWOKRQKQVNDENNVMNVEYSSELEKHQL-----YIDETVNSNIATNLR	196						
DB	93	YNPOQPKPGMIDSATQSKWVEE-----IVKYEALLLTHETSRYLQEIYNSN-----	142						
QY	197	VLRSIENLRSKIQKLESVDYSAOMEYCRTPCTVSCNIPVWSGKECEBIIIRKGGTSEMYL	256						
DB	143	--NQKITNLKQVLALEAQ-----CQEPCKDSVQIHDITGKQCBIANKGAKGESGIYF	193						
QY	257	IOPDSSVKPYRYVYCDMNTENGWTVIQRQDGSVDPRKRWDPYKQGGGNVA-TWTDGKNY	315						
DB	194	IRPLKAKQFLVYCEIDSGNGWTVLQKRIDGSLDFKKNWITQYKEGFGHLSPGTGT-----	249						
QY	316	GLPGEYWLGNDKISQLFRMG-----PTELLIEMEDWKGDKVKAHYGGFTVQNEANKYQISVN	373						
DB	250	-----EFWLGNKEKHLISMQSTIPIYALRIQLDNWGRSTADYAMFRVGPESDKRYLTYA	304						
QY	374	KY-RGTAGNALMDGASQLMGE---NRTWTIHNGMFFSTYDRDNDGWLTS DPRKQCSKEDG	429						
DB	305	YFIGDGADGA-FDGYD--FGDDPSDKFTTSINGMQFTSWDNDNDKF-----EGNCAEQDG	356						
QY	430	GGWYNYNRCHAANPNRYTWGGGYTWDMAKHGTTDGVVWNNWKGWYSMRKMSMKIRPF	487						

```

Db          357 SCWNNKCHAGHLNGVHHGGTYSKSSSTNGF
search completed: November 22, 2004, 18:39:40
Job time : 199 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 22, 2004, 15:54:56 ; Search time 14396.5 Seconds
(without alignments)
12349.531 Million cell updates/sec

Title: US-10-017-724-5_COPY_4000_8878

Perfect score: 4879

Sequence: 1 tcagaatagtgtgtgatgag.....gattgattgagccaagatc 4879

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 32822875 seqs, 18219865908 residues

Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : EST:*

1: gb_est1: *
2: gb_est2: *
3: gb_hic: *
4: gb_est3: *
5: gb_est4: *
6: gb_est5: *
7: gb_est6: *
8: gb_gss1: *
9: gb_gss2: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	696.6	14.3	716	5	BX118749
2	686	14.1	768	4	BG563877
3	671.6	13.8	820	7	CO775659
4	662.4	13.6	758	4	BG568684
5	655.2	13.4	679	4	BM683513
6	650.4	13.3	673	8	AQ489171
7	644.4	13.2	1891	3	CR594769
8	643.4	13.2	1577	3	CR596720
9	642.4	13.2	1883	3	CR596576
10	641.4	13.1	1888	3	CR611717
11	640.4	13.1	1014	5	BX448857
12	639.4	13.1	1886	3	CR596719
13	637.8	13.1	997	5	BX427240
14	634.4	13.0	1880	3	CR626133
15	634	13.0	804	1	AW003933
16	634	13.0	1016	5	BX427173
17	632.4	13.0	1882	3	CR593920
18	628.4	12.9	1038	5	BX448675
19	628.4	12.9	1577	3	CR615740
20	628.4	12.9	1875	3	CR614263
21	627.4	12.9	1873	3	CR609458
22	626.2	12.8	1038	5	BX405660
23	625.4	12.8	1872	3	CR604009
24	624.4	12.8	1869	3	CR617322

C 25	624	12.8	937	5	BX346072	BX346072
C 26	624	12.8	1041	5	BX346089	BX346089
C 27	623.4	12.8	1572	3	CR598319	CR598319
C 28	623.4	12.8	1848	3	CR597172	full-length
C 29	622.6	12.8	1064	5	BX427236	BX427236
C 30	621.4	12.7	1862	3	CR594221	full-length
C 31	621.2	12.7	1045	5	BX448329	BX448329
C 32	618	12.7	930	1	AL564576	AL564576
C 33	616.4	12.6	1859	3	CR602828	full-length
C 34	615.6	12.6	699	1	AV652183	AV652183
C 35	615.4	12.6	636	2	AW771537	AW771537
C 36	615	12.6	1007	1	AL579878	AL579878
C 37	614.2	12.6	936	5	BX423808	BX423808
C 38	611.4	12.5	1029	1	AL579779	AL579779
C 39	609.8	12.5	1008	5	BX448288	BX448288
C 40	608.2	12.5	664	1	AV681947	AV681947
C 41	608.2	12.5	1027	5	BX405881	BX405881
C 42	607.6	12.5	628	5	BX494484	BX494484
C 43	606.2	12.4	1029	5	BX427659	BX427659
C 44	604.8	12.4	761	1	AV700652	AV700652
C 45	604.6	12.4	1009	1	AL531649	AL531649

ALIGNMENTS

RESULT 1
BX118749/c
LOCUS BX118749 716 bp mRNA linear EST 10-FEB-2003
DEFINITION IMAGp998P03372 ; IMAGE:195698, mRNA sequence.
ACCESSION BX118749
VERSION BX118749.1 GI:27841363
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 716)
AUTHORS Ebert,L., Heil,O., Hennig,S., Neubert,P., Partsch,E., Peters,M., Radloff,U., Schneider,D. and Korn,B.
TITLE Human Unigeneset - RZPD3
JOURNAL Unpublished (2003)
COMMENT Contact: Ina Rolfs
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
Im Neuenheimer Feld 580, D-69120 Heidelberg, Germany
RZPD; IMAGp998P03372.
RZPDLIB: I.M.A.G.E. cDNA Clone Collection;
http://www.rzpd.de/CloneCards/cgi-bin/showLib.pl.cgi?response?libNo=972 Contact: Ina Rolfs
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
Heubnerweg 6, D-14059 Berlin, Germany
Tel: +49 30 32639 101
Fax: +49 30 32639 111
www.rzpd.de
This clone is available royalty-free from RZPD;
contact RZPD (clone@rzpd.de) for further information. Seq primer:
M13r, Primer sequence: TTTACACAGGAAACAGCTATGAC.

FEATURES

Location/Qualifiers
1..716
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGp998P03372 ; IMAGE:195698"
/sex="male"
/dev_stage="20 week-post conception fetus"
/lab_host="DH10B (ampicillin resistant)"
/clone_lib="Soares fetal liver spleen INFLS"
/note="Organ: Liver and spleen; Vector: pT73D (Pharmacia) with a modified polylinker; Site 1: Pac I; Site 2: Eco RI; 1st strand cDNA was primed with a Pac I - oligo(dT) primer [5' ACTGGAAGATTATTAAGAATCTTTTTTTTTTTTTTTT 3']

Query Match	13.8%;	Score 671.6;	DB 7;	Length 820;
Best Local Similarity	91.6%;	Pred. No. 6.6e-106;		
Matches 744;	Conservative 0;	Mismatches 64;	Indels 4;	Gaps 3;
QY	3672	TTGCCACAAGGCCAGTTATCTCTCTTTGCTATAGGCCACAGGAGGCTTTTGGTGT	3731	

RESULT 4	ACCESSION	REFERENCE
BG568684	VERSION	AUTHORS
LOCUS	KEYWORDS	TITLE
DEFINITION	SOURCE	JOURNAL
	ORGANISM	COMMENT

QY 2947 TTTATTTTGTCTTTTCTTTAGTGAATATTTGGCTTGGAAATGATAAAATAGCCAGCTTA 3006
 Db 540 TTTATTTTGTATTTCTTTTGTAGTGAATATTTGGCTTGGAAATGATAAAATAGCCAGCTTA 599
 QY 3007 CCAGATGGGACCCACAGAACTTTTGTATAGAAATGGAGGACTGGAAGGAGACAAAGTAA 3066
 Db 600 CCAGATGGGACCCACAGAACTTTTGTATAGAAATGGAGGACTGGAAGGAGACAAAGTAA 659
 QY 3067 AGGCTCACTATGGA 3080
 Db 660 AGGCTCACTATGGA 673
 RESULT 7
 CR594769 1891 bp mRNA linear HTC 21-JUL-2004
 LOCUS full-length cDNA clone CS0DM011Y115 of Fetal liver of Homo sapiens (human).
 DEFINITION
 ACCESSION CR594769
 VERSION 1 GI:50475576
 KEYWORDS HTC; CNSLT_cDNA.
 SOURCE Homo sapiens
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 1891)
 AUTHORS Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
 TITLE Full-length cDNA libraries and normalization
 JOURNAL Unpublished
 REMARK Contact : Feng Liang Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com/ InvitroGen Corporation 1600 Faraday Avenue
 REFERENCE 2 (bases 1 to 1891)
 AUTHORS Genoscope.
 TITLE Direct Submission
 JOURNAL Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)
 COMMENT 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies, a division of Invitrogen.
 FEATURES
 Location/Qualifiers
 1..1891
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CS0DM011Y115"
 /tissue_type="Fetal liver"
 /plasmid="pCMVSPORT_6"
 ORIGIN
 Query Match 13.2%; Score 644.4; DB 3; Length 1891;
 Best Local Similarity 99.7%; Pred. No. 2,9e-101;
 Matches 656; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
 QY 3869 CAGGTATACATCAGATCCAGAAAACAGTGTCTTAAGAGACGGTGGTGGATGGTGA 3928
 Db 1235 CTGGTTAATCATCAGATCCAGAAAACAGTGTCTTAAGAGACGGTGGTGGATGGTGA 1294
 QY 3929 TAATAGATGTCATGACGCAATCCAAACGGCAGATCTACTGGGTGGACAGTACACCTG 3988
 Db 1295 TAATAGATGTCATGACGCAATCCAAACGGCAGATCTACTGGGTGGACAGTACACCTG 1354
 QY 3989 GGACATGGCAAGACATGGCACAGATGATGGTGTAGTATGGATGAATGGAAGGGTCAATG 4048
 Db 1355 GGACATGGCAAGACATGGCACAGATGATGGTGTAGTATGGATGAATGGAAGGGTCAATG 1414
 QY 4049 GTACTCAATGAGGAGATGAGTATGAGATCAGGCCCTTCTCCACACCAATAGTCCCC 4108
 Db 1415 GTACTCAATGAGGAGATGAGTATGAGATCAGGCCCTTCTCCACACCAATAGTCCCC 1474

QY 4109 AATACGTAGATTTTGTCTCTTCTGTATGTGACAACTTTTGTGATCATATGTTATGGAA 4168
 Db 1475 AATACGTAGATTTTGTCTCTTCTGTATGTGACAACTTTTGTGATCATATGTTATGGAA 1534
 QY 4169 TTTCTTTTCATACATATATATCTCTTAAACTCTCAAGCAGACGCTGAGTGTGACTTTTGTG 4228
 Db 1535 TTTCTTTTCATACATATATATCTCTTAAACTCTCAAGCAGACGCTGAGTGTGACTTTTGTG 1594
 QY 4229 AAAAAAGTATAGGATAAATTACATTTAAATATAGCATGATTTTCTTTTGTCTTCTCAT 4288
 Db 1595 AAAAAAGTATAGGATAAATTACATTTAAATATAGCATGATTTTCTTTTGTCTTCTCAT 1654
 QY 4289 TCTTCTCTCACCACGAAGTAAACAAAGTATATAGTTTTCAGAGATGGTGGTTCATTAAT 4348
 Db 1655 TCTTCTCTCA-CCAAGAAAGTAAACAAAGTATATAGTTTTCAGAGATGGTGGTTCATTAAT 1713
 QY 4349 TCAGTTCTAGTTGATTCGGAGAAATTTCAATATAGGAGAGGGGTCTTTTATCTTCTGTCG 4408
 Db 1714 TCAGTTCTAGTTGATTCGGAGAAATTTCAATATAGGAGAGGGGTCTTTTATCTTCTGTCG 1773
 QY 4409 TAGGAAAACCATGACGGAAAGGAAAACTGATGTTTAAAGTCCACTTTTAAACTATAT 4468
 Db 1774 TAGGAAAACCATGACGGAAAGGAAAACTGATGTTTAAAGTCCACTTTTAAACTATAT 1833
 QY 4469 TTATTTATGTAGGATCTGTCAAAGAAAACCTTCCAAAAGAGATTATTAATTAACACCA 4526
 Db 1834 TTAUTTATGTAGGATCTGTCAAAGAAAACCTTCCAAAAGAGATTATTAATTAACACCA 1891
 RESULT 8
 CR596720 1577 bp mRNA linear HTC 21-JUL-2004
 LOCUS full-length cDNA clone CS0DM009Y104 of Fetal liver of Homo sapiens (human).
 DEFINITION
 ACCESSION CR596720
 VERSION 1 GI:50477527
 KEYWORDS HTC; CNSLT_cDNA.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 1577)
 AUTHORS Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
 TITLE Full-length cDNA libraries and normalization
 JOURNAL Unpublished
 REMARK Contact : Feng Liang Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com/ InvitroGen Corporation 1600 Faraday Avenue
 REFERENCE 2 (bases 1 to 1577)
 AUTHORS Genoscope.
 TITLE Direct Submission
 JOURNAL Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)
 COMMENT 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies, a division of Invitrogen.
 FEATURES
 Location/Qualifiers
 1..1577
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CS0DM009Y104"
 /tissue_type="Fetal liver"
 /plasmid="pCMVSPORT_6"
 ORIGIN
 Query Match 13.2%; Score 643.4; DB 3; Length 1577;
 Best Local Similarity 99.7%; Pred. No. 4,5e-101;
 Matches 655; Conservative 0; Mismatches 1; Indels 1; Gaps 1;


```

QY 4169 TTTTCTTTCATATATATTCCTTAAACTCTCAAGCAGACGTGAGTGTGACTTTTGG 4228
Db 355 TTTTCTTTCATATATATTCCTTAAACTCTCAAGCAGACGTGAGTGTGACTTTTGG 296
QY 4229 AAAAAAGTAGAGATAAATTACATTAATAAATAGCAGATGATTTCTTTGTTTCTTCATT 4288
Db 295 AAAAAAGTAGAGATAAATTACATTAATAAATAGCAGATGATTTCTTTGTTTCTTCATT 236
QY 4289 TCTCTTGTCTCACCACCAAGAGTAACAAAGTATATGTTTTCACAGAGTGTGTTTCATAATT 4348
Db 235 TCTCTTGTCTCA - CCAAGAGTATACAAAGTATATGTTTTCACAGAGTGTGTTTCATAATT 177
QY 4349 TCAGTTCTAGTTGATTCGAGAAATTTTCAATAAAGGAAGAGGGGTCTTTTATCCTTGTGCG 4408
Db 176 TCAGTTCTAGTTGATTCGAGAAATTTTCAATAAAGGAAGAGGGGTCTTTTATCCTTGTGCG 117
QY 4409 TAGGAACCATGACGGAAGGAAGAACTGATGTTTAAAGTCCACTTTTAAACTATAT 4468
Db 116 TAGGAACCATGACGGAAGGAAGAACTGATGTTTAAAGTCCACTTTTAAACTATAT 57
QY 4469 TTATTTATGATGATCTGTCAGAAAGAACTTCCAAAGAGATTTATTAATAACCA 4524
Db 56 TTATTTATGATGATCTGTCAGAAAGAACTTCCAAAGAGATTTATTAATAACCA 1

RESULT 12
LOCUS CR596719
DEFINITION full-length cDNA clone CS0DM009Y103 of Fetal liver of Homo sapiens
(human).
ACCESSION CR596719
VERSION CR596719.1 GI:50477526
KEYWORDS HTC; CNSLT cDNA.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1886)
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished
Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue
Genoscope,
2 (bases 1 to 1886)
Direct Submission
Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
FEATURES
Location/Qualifiers
1..1886
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DM009Y103"
/tissue_type="Fetal liver"
/plasmid="pCMVSPORT_6"

ORIGIN
Query Match 13.1%; Score 639.4; DB 3; Length 1886;
Best Local Similarity 99.7%; Pred. No. 2.2e-100;
Matches 651; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 3869 CAGGTTACATCAGATCCAGAAAACAGTGTCTAAAGAACGCGTGTGGATGGTGA 3928
Db 1235 CTGGTTAAACATCAGATCCAGAAAACAGTGTCTAAAGAACGCGTGTGGATGGTGA 1294

```

```

QY 3929 TAATAGATGTCATGCGAGCCAAATCCAAACGCGAGATACTACTGGGGTGGACAGTACACCTG 3988
Db 1295 TAATAGATGTCATGCGAGCCAAATCCAAACGCGAGATACTACTGGGGTGGACAGTACACCTG 1354
QY 3989 GGACATGCGCAAGCAGTGGCAGACAGATGATGTTGATGATCGATGAATTTGGAAGGGGTCAATG 4048
Db 1355 GGACATGCGCAAGCAGTGGCAGACAGATGATGTTGATGATCGATGAATTTGGAAGGGGTCAATG 1414
QY 4049 GTACTCAATGAGGAAGATGAGTATGAAGATCAGGCCCTTTCTTCCACAGCAATAGTCCCTCC 4108
Db 1415 GTACTCAATGAGGAAGATGAGTATGAAGATCAGGCCCTTTCTTCCACAGCAATAGTCCCTCC 1474
QY 4109 AATAGCTAGATTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 4168
Db 1475 AATAGCTAGATTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1534
QY 4169 TTTTCTTTCATACATATATATTCCTTAAACTCTCAAGCAGACGTGAGTGTGACTTTTGG 4228
Db 1535 TTTTCTTTCATACATATATATTCCTTAAACTCTCAAGCAGACGTGAGTGTGACTTTTGG 1594
QY 4229 AAAAAAGTAGAGATAAATTACATTAATAAATAGCAGATGATTTCTTTGTTTCTTCATT 4288
Db 1595 AAAAAAGTAGAGATAAATTACATTAATAAATAGCAGATGATTTCTTTGTTTCTTCATT 1654
QY 4289 TCTCTTGTCTCACCACCAAGAGTAACAAAGTATATGTTTTCACAGAGTGTGTTTCATAATT 4348
Db 1655 TCTCTTGTCTCA - CCAAGAGTATACAAAGTATATGTTTTCACAGAGTGTGTTTCATAATT 1713
QY 4349 TCAGTTCTAGTTGATTCGAGAAATTTTCAATAAAGGAAGAGGGGTCTTTTATCCTTGTGCG 4408
Db 1714 TCAGTTCTAGTTGATTCGAGAAATTTTCAATAAAGGAAGAGGGGTCTTTTATCCTTGTGCG 1773
QY 4409 TAGGAACCATGACGGAAGGAAGAACTGATGTTTAAAGTCCACTTTTAAACTATAT 4468
Db 1774 TAGGAACCATGACGGAAGGAAGAACTGATGTTTAAAGTCCACTTTTAAACTATAT 1833
QY 4469 TTATTTATGATGATCTGTCAGAAAGAACTTCCAAAGAGATTTATTAATAACCA 4521
Db 1834 TTATTTATGATGATCTGTCAGAAAGAACTTCCAAAGAGATTTATTAATAACCA 1886

RESULT 13
LOCUS BX427240/c
DEFINITION BX427240 Homo sapiens FETAL LIVER Homo sapiens cDNA clone
CS0DM009Y104 3-PRIME, mRNA sequence.
ACCESSION BX427240
VERSION BX427240.1 GI:30772645
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 997)
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
Contact : Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library
was not normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
This sequence belongs to sequence cluster 2627.f
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?e=CS0AM009BE02NP1&c=2627.f.
FEATURES
Location/Qualifiers
1..997
/organism="Homo sapiens"
source

```

```
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DM009Y104"
/tissue_type="FETAL LIVER"
/dev_stage="fetal"
/clone_lib="Homo sapiens FETAL LIVER"
/notes="Organ: liver; Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized."
```

ORIGIN

Query Match 13.1%; Score 637.8; DB 5; Length 997;
Best Local Similarity 98.3%; Pred. No. 4.5e-100;
Matches 646; Conservative 7; Mismatches 3; Indels 1; Gaps 1;

QY 3869 CAGGTTAAACATCAGATCCAGAAACACAGTGTCTTAAAGAACGCGTGTGATGGTGA 3928
DB 656 CTGGTTAAACATCAGATCCWAAAAACAGTGTCTTAAAGAACGCGTGTGATGGTGA 597

QY 3929 TAATAGATGTCTATGCGAGCAATCCAAACGCGAGATCTACTGGGTGGACAGTACACCTG 3988
DB 596 TAATAGATGTCTATGCGAGCAATCCAAACGCGAGATCTACTGGGTGGACAGTACACCTG 537

QY 3989 GGACATGCCAAAGCATGCGACAGATGATGGTGTAGTATGGATGAATTCGAAGGGGTCTATG 4048
DB 536 GGACATGCCAAAGCATGCGACAGATGATGGTGTAGTATGGATGAATTCGAAGGGGTCTATG 477

QY 4049 GTACTCAATAGGAAGATGAGTATGAAGATCAGGCCCTTCTTCCACAGCAATAGTCCCC 4108
DB 476 GTACTCAATAGGAAGATGAGTATGAAGATCAGGCCCTTCTTCCACAGCAATAGTCCCC 417

QY 4109 AATAGTAGATTTTGTCTTCTCTGTATGTGACAAATTTTGTACATTTATTTGGAA 4168
DB 416 AATAGTAGATTTTGTCTTCTCTGTATGTGACAAATTTTGTACATTTATTTGGAA 357

QY 4169 TTTTCTTTCATACATTTATTTCTCTTAAACCTCTCAAGCAGACGCTGAGTGTGACATTTTGTG 4228
DB 356 TTTTCTTTCATACATTTATTTCTCTTAAACCTCTCAAGCAGACGCTGAGTGTGACATTTTGTG 297

QY 4229 AAAAAAGTAGGATAAATTAACATTAATAAGCAGATGATTTCTTTTGTCTTCTCAT 4288
DB 296 AAAAAAGTAGGATAAATTAACATTAATAAGCAGATGATTTCTTTTGTCTTCTCAT 237

QY 4289 TCTCTTGTCTACCCGAAGATTAACAAAGTATAGTTTTCAGACAGCTGGTGTTCATAAT 4348
DB 236 TCTCTTGTCTCA-CCAAGAAAGTAAACAAAGTATAGTTTTCAGACAGCTGGTGTTCATAAT 178

QY 4349 TCAGTTCTAGTTGATGCGAGAAATTTTCAATAAGGAAGAGGGGTCTTTTATCCTTGTGCG 4408
DB 177 TCAGTTCTAGTTGATGCGAGAAATTTTCAATAAGGAAGAGGGGTCTTTTATCCTTGTGCG 118

QY 4409 TAGGAAAACCATGACGGAAAGAAACATGTGTTTAAAGTCCACTTTTAAACATATAT 4468
DB 117 TAGGAAAACCATGACGGAAAGAAACATGTGTTTAAAGTCCACTTTTAAACATATAT 58

QY 4469 TTATTATGATGATCTGTCAAGAAACCTTCCAAAGAGATTTATTAATTAACAG 4525
DB 57 TTATTATGATGATCTGTCAAGAAACCTTCCAAAGAGATTTVKNKAADWAACCCAG 1

RESULT 14
CR626133
LOCUS
DEFINITION
full-length cDNA clone CS0DH008Y104 of T cells (Jurkat cell line)
of Homo sapiens (human).
CR626133
ACCESSION
CR626133.1 GI:50506940
VERSION
HMC; CNSLT_cDNA.
KEYWORDS
Homo sapiens (human)
Homo sapiens
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1880)
Li, W.B., Gruber, C., Jesse, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished
Contact: Feng Liang Email: fliang@lifetech.com URL: <http://fulllength.invitrogen.com/> Invitrogen Corporation 1600
Faraday Avenue
2 (bases 1 to 1880)
Genoscope.
Direct Submission
Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail: secref@genoscope.cns.fr)
- Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
Location/Qualifiers
1. 1880
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DH008Y104"
/tissue_type="T cells (Jurkat cell line)"
/plasmid="pCMVSPORT_6"

ORIGIN

Query Match 13.0%; Score 634.4; DB 3; Length 1880;
Best Local Similarity 99.7%; Pred. No. 1.6e-99;
Matches 646; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 3869 CAGGTTAAACATCAGATCCAGAAACACAGTGTCTTAAAGAACGCGTGTGATGGTGA 3928
DB 1234 CTGGTTAAACATCAGATCCAGAAACACAGTGTCTTAAAGAACGCGTGTGATGGTGA 1293

QY 3929 TAATAGATGTCTATGCGAGCAATCCAAACGCGAGATCTACTGGGTGGACAGTACACCTG 3988
DB 1294 TAATAGATGTCTATGCGAGCAATCCAAACGCGAGATCTACTGGGTGGACAGTACACCTG 1353

QY 3989 GGACATGCCAAAGCATGCGACAGATGATGGTGTAGTATGGATGAATTCGAAGGGGTCTATG 4048
DB 1354 GGACATGCCAAAGCATGCGACAGATGATGGTGTAGTATGGATGAATTCGAAGGGGTCTATG 1413

QY 4049 GTACTCAATAGGAAGATGAGTATGAAGATCAGGCCCTTCTTCCACAGCAATAGTCCCC 4108
DB 1414 GTACTCAATAGGAAGATGAGTATGAAGATCAGGCCCTTCTTCCACAGCAATAGTCCCC 1473

QY 4109 AATAGTAGATTTTGTCTTCTCTGTATGTGACAAATTTTGTGACATTTATTTGGAA 4168
DB 1474 AATAGTAGATTTTGTCTTCTCTGTATGTGACAAATTTTGTGACATTTATTTGGAA 1533

QY 4169 TTTTCTTTCATACATTTATTTCTCTTAAACCTCTCAAGCAGACGCTGAGTGTGACATTTTGTG 4228
DB 1534 TTTTCTTTCATACATTTATTTCTCTTAAACCTCTCAAGCAGACGCTGAGTGTGACATTTTGTG 1593

QY 4229 AAAAAAGTAGGATAAATTAACATTAATAAGCAGATGATTTCTTTTGTCTTCTCAT 4288
DB 1594 AAAAAAGTAGGATAAATTAACATTAATAAGCAGATGATTTCTTTTGTCTTCTCAT 1653

QY 4289 TCTCTTGTCTACCCGAAGATTAACAAAGTATAGTTTTCAGACAGCTGGTGTTCATAAT 4348
DB 1654 TCTCTTGTCTCA-CCAAGAAAGTAAACAAAGTATAGTTTTCAGACAGCTGGTGTTCATAAT 1712

QY 4349 TCAGTTCTAGTTGATGCGAGAAATTTTCAATAAGGAAGAGGGGTCTTTTATCCTTGTGCG 4408
DB 1713 TCAGTTCTAGTTGATGCGAGAAATTTTCAATAAGGAAGAGGGGTCTTTTATCCTTGTGCG 1772

QY 4409 TAGGAAAACCATGACGGAAAGAAACATGTGTTTAAAGTCCACTTTTAAACATATAT 4468
DB 1773 TAGGAAAACCATGACGGAAAGAAACATGTGTTTAAAGTCCACTTTTAAACATATAT 1832

QY 4469 TTATTATAGGATCTGTCAAGAAACTTCCAAAAGATTATTAA 4516
Db 1833 TTATTATAGGATCTGTCAAGAAACTTCCAAAAGATTATTAA 1880

RESULT 15
AW003933/c
LOCUS
DEFINITION
q84e02.x1 NCI CGAP GC6 Homo sapiens cDNA clone IMAGE:2478746 3,
similar to gb:M64983_rna6 FIBRINOGEN BETA CHAIN PRECURSOR (HUMAN);,
mRNA sequence.

ACCESSION
AW003933
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)

REFERENCE
1 (bases 1 to 804)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

AUTHORS
TITLE
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)

JOURNAL
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov
Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael
R. Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
Bonaldo, Ph.D.

FEATURES
source
1..804
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:2478746"
/tissue_type="pooled germ cell tumors"
/lab_host="DH10B"
/clone_lib="NCI CGAP GC6"
/note="Vector: pT7T3D-Pac (Pharmacia) with a modified
polylinker; Site1: Not 1; Site 2: Eco RI; Plasmid DNA
from the normalized library NCI CGAP GC4 was prepared, and
ss circles were made in vitro. Following HAP purification,
this DNA was used as tracer in a subtractive hybridization
reaction. The driver was PCR-amplified cDNAs from a pool
of 5,000 clones made from the same library (cloneIDs
1257096-1258631, 1469064-1470983, and 1475592-1476743).
Subtraction by Bento Soares and M. Fatima Bonaldo."

ORIGIN
Query Match 13.0%; Score 634; DB 1; Length 804;
Best Local Similarity 97.6%; Pred. No. 2.1e-99;
Matches 654; Conservative 0; Mismatches 15; Indels 1; Gaps 1;

QY 3869 CAGGTTAAACATCAGATCCAGAAACAGTGTCTAAAGAACGCGTGGTGGTGGTGA 3928
Db 669 CTGGTTAAACATCAGATCCAGAAACAGTGTCTAAAGAACGCGTGGTGGTGGTGA 610

QY 3929 TAATAGATGTCATGACGCAATCCAAACGGCAGATACCTAGTGGGTGGACAGTACACCTG 3988
Db 609 TAATAGATGTCATGACGCAATCCCAATCCATACCTAGTGGGTGGACAGTACACCTG 550

QY 3989 GGACATGGCAAGCATGGCAGAGATGCGTGTAGTATGCGAATTCGAAGGGGTGATG 4048
Db 549 GGACATGGCAAGCATGGCAGAGATGCGTGTAGTATGCGAATTCGAAGGGGTGATG 490

QY 4049 GTACTCTAAGGAAGATGAGTATGAAGATCAGGGCCCTTCTCCACAGCAATAGTCCCC 4108

Db 489 GTACTCAATGAAGAGATGAGTATGAAGATCAGGCCCTTCTCCACAGCAATAGTCCCC 430
QY 4109 AATAGTAGATTTTGTCTCTCTGTATGTGACAAACATTTTGTACATTAATGTTATGGAA 4168
Db 429 AATAGTAGATTTTGTCTCTCTGTATGTGACAAACATTTTGTACATTAATGTTATGGAA 370
QY 4169 TTTTCTTTTCATACATTAATTTCTCTAAACTCTCAAGCAGACGTCAGTGTGACTTTTGG 4228
Db 369 TTTTCTTTTCATACATTAATTTCTCTAAACTCTCAAGCAGACGTCAGTGTGACTTTTGG 310
QY 4229 AAAAAAGTATAGGATAAATTACATTAATAATAGCACAATGATTTTCTTTTCTTTCTTCAT 4288
Db 309 AACAAAGTATAGGATCAATTACATTAACATAGCACAATGATTTTCTTTTCTTTCTTCAT 250
QY 4289 TCTCTTGTCTCACCCAGAAAGTAAACAAAGTATAGTTTGTGACAGAGTGGTGTTCATAAT 4348
Db 249 TCTCTTGTCTCA-CCAAAGAAAGTAAACAAAGTATAGTTTGTGACAGAGTGGTGTTCATAAT 191
QY 4349 TCAGTTCTAGTTGATTCGAGAAATTTCAAAATAAGGAAGAGGGGTCTTTTATCCTTGTG 4408
Db 190 TCAGTTCTAGTTGATTCGAGAAATTTCAAAATAAGGAAGAGGGGTCTTTTATCCTTGTG 131
QY 4409 TAGGAAAAACCATGACGAAAGGAAAACTGATGTTTAAAGTCCACTTTTAAACCTATAT 4468
Db 130 TAGGAAAAACCATGACGAAAGGAAAACTGATGTTTAAAGTCCACTTTTAAACCTATAT 71
QY 4469 TTATTTATGTAGGATCTGTCAAGAAAACTTCCAAAAAGATTTATTAATTAACACGACT 4528
Db 70 TTATTTATGTAGGATCTGTCAAGAAAACTTCCAAAAAGATTTATTAATTAACACGACT 11
QY 4529 CTGTTGCAAT 4538
Db 10 CTGTTGCAAT 1

Search completed: November 23, 2004, 12:26:29
Job time : 14403.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 22, 2004, 18:22:12 ; Search time 155 Seconds
(without alignments)
1136.362 Million cell updates/sec

Title: US-10-017-724-6
Perfect score: 2676
Sequence: 1 MKRMVSWFPHKLTWKHLL.....GSWYSMRKSMKIRPFPPQQ 491

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_23Sep04:*

- 1: Geneseqp1980s:*
- 2: Geneseqp1990s:*
- 3: Geneseqp2000s:*
- 4: Geneseqp2001s:*
- 5: Geneseqp2002s:*
- 6: Geneseqp2003as:*
- 7: Geneseqp2003bs:*
- 8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2676	100.0	491	6	Aae36413 Human FBG
2	2676	100.0	491	7	Adp65229 Human fib
3	2676	100.0	491	7	Adp65300 Human fib
4	2671	99.8	491	2	Aar82243 Human fib
5	2671	99.8	491	4	Aam78492 Human pro
6	2651	99.1	495	4	Aam78492 Human pro
7	2651	99.1	495	4	Aam79475 Human pro
8	2651	99.1	495	4	Aam79477 Human pro
9	2648	99.0	488	8	Adp65229 Human pro
10	2637	98.5	539	4	Aam78491 Human pro
11	2559	95.6	491	7	Adp65229 Human fib
12	2559	95.6	491	7	Adp65229 Human fib
13	2436	91.0	453	4	Aam78493 Human pro
14	2170	81.1	479	8	Adi82107 Human pro
15	1072.5	40.1	489	5	Abu60918 Lung spec
16	1001	37.4	502	4	Aau31380 Novel hum
17	785	29.3	150	3	Aag00150 Human sec
18	630.5	23.6	437	8	Adk70498 Respirato
19	630.5	23.6	437	8	Adl61247 Human pro
20	630.5	23.6	453	7	Adl61247 Human pro
21	630.5	23.6	453	8	Adc45364 Human pro
22	629.5	23.5	453	8	Adc30579 Pancreas
23	629.5	23.5	453	2	Aar82246 Human fib
24	620.5	23.2	411	2	Aar82245 Human fib
25	610	22.8	141	4	Aar84650 Fibrinogen
				4	Aao10778 Human pol

ALIGNMENTS

RESULT 1
ID AAE36413 standard; protein; 491 AA.
XX AAE36413;
AC
XX

DT 07-AUG-2003 (first entry)
XX

DE Human FBG reference protein (GI 11761631).
XX

KW Thrombospondin 2; THBS2; angiotensin converting enzyme; polymorphism;
ACE-1; beta-fibrinogen; FGB; peripheral vascular disease; ischaemia;
KW vascular disease; myocardial infarction; pulmonary embolism; stroke;
KW atherosclerosis; coronary artery disease; venous thromboembolism; human.
XX Homo sapiens.
XX

PH Key Location/Qualifiers
FT Misc-difference 478
FT /note= "This residue is changed to Lys due to single
FT nucleotide polymorphism (SNP)"
XX

PN WO2003020118-A2.
XX

PD 13-MAR-2003.
XX

PF 04-SEP-2002; 2002WO-US028113.
XX

PR 05-SEP-2001; 2001US-0317178P.
PR 16-OCT-2001; 2001US-0329958P.
PR 14-DEC-2001; 2001US-00017724.
XX

PA (VITI-) VITIVITY INC.
XX

PI McCarthy J;
XX

DR WPI; 2003-300816/29.
XX

PT Identifying polymorphisms in thrombospondin 2, angiotensin converting
enzyme and/or beta-fibrinogen genes in nucleic acid sample of subject, by
contacting the nucleic acid with a complementary probe or primer.
XX Claim 5; Fig 6; 194pp; English.

XX The invention relates to a method for determining the identity of one or
more allelic variants of a polymorphic region of a thrombospondin 2
(THBS2), angiotensin converting enzyme (ACE)-1 and/or beta-fibrinogen
(FGB) genes in a nucleic acid obtained from a subject. The method

301 QGFGNVAITDGNKYCGLPGEYWLNDKISQLTRMGPTTELLIEMEDWKGDKVKAHYGGFT 360
301 QGFGNVAITDGNKYCGLPGEYWLNDKISQLTRMGPTTELLIEMEDWKGDKVKAHYGGFT 360
361 VQNEANKYQISVKNYKGTAGNALMDGASQIMGNRTWTTHNGMFFSTYDRDNDGWLTSDD 420
361 VQNEANKYQISVKNYKGTAGNALMDGASQIMGNRTWTTHNGMFFSTYDRDNDGWLTSDD 420
421 RQCKSKEDGGGWWYNNRCHAANPNRYWGQYTWDMAXHGTDDGVVMMNWKGSWYSMRKM 480
421 RQCKSKEDGGGWWYNNRCHAANPNRYWGQYTWDMAXHGTDDGVVMMNWKGSWYSMRKM 480
481 SMKIRPPFPQQ 491
481 SMKIRPPFPQQ 491

RESULT 3
ADP65300
ID ADP65300 standard; protein; 491 AA.
AC ADP65300;
DT 12-AUG-2004 (first entry)
DE Human fibrinogen, beta chain preproprotein.
KW autoimmune disease; arthritis; gene expression analysis;
KW rheumatoid arthritis; collagen-induced; immunosuppressive; antirheumatic;
KW antiarthritic; osteopathic; angiot; antinflammatory; dermatological;
KW immunomodulatory; lupus; ankylosing spondylitis; Fibrositis;
KW fibromyalgia; osteoarthritis; gout; juvenile rheumatoid arthritis;
KW immune; human.
OS Homo sapiens.
PN WO2003072827-A1.
XX
PD 04-SEP-2003.
PF 31-OCT-2002; 2002WO-US035433.
PR 31-OCT-2001; 2001US-0336220P.
PA (CHIL-) CHILDREN'S HOSPITAL MEDICAL CENT.
PI Hirsch R, Thorton SL;
PI WPI; 2003-712740/67.
DR GENBANK; NP_005132.
XX
PT Diagnosing and analyzing autoimmune disease using gene expression
PT profiles and microarray technology, useful for diagnosing and treating
PT rheumatoid arthritis, lupus, fibrositis, osteoarthritis, fibromyalgia and
PT gout.
PS Disclosure: Page; 56pp; English.

The invention relates to a novel method for diagnosing and analysing
autoimmune disease or arthritides. The method comprises obtaining a
patient sample containing mRNA, analysing gene expression using the mRNA
that results in a gene expression signature of the mRNA, and using that
gene expression signature to diagnose or analyse the autoimmune disease
or arthritides in the patient, where gene expression of at least 60% of
the genes correlates with that of the gene signature. The invention
further comprises: a treatment of rheumatoid arthritis; identification of
genes for targeting in the treatment of rheumatoid arthritis in a mammal
other than a mouse; diagnosis of rheumatoid arthritis in a mammal
array or gene chip, specific for rheumatoid arthritis; diagnosis or
analyses of autoimmune disease or rheumatoid arthritis; screening the
efficacy of a candidate drug in vitro for the treatment of collagen-
induced arthritis; and reducing the symptoms associated with collagen-
induced arthritis. The compositions of the invention have the following

CC activities: immunosuppressive, antirheumatic, antiarthritic, osteopathic,
CC angiot, antinflammatory, dermatological, and immunomodulatory. The
CC methods and compositions of the present invention are useful for
CC diagnosing and treating autoimmune disease or arthritides, such as
CC rheumatoid arthritis, lupus, ankylosing spondylitis, fibrositis,
CC fibromyalgia, osteoarthritis, gout, juvenile rheumatoid arthritis, and an
CC immune disease caused by an infectious agent. This sequence represents a
CC protein sequence relating to the genes used in the analysis and treatment
CC of autoimmune diseases or arthritides. Note: This sequence is not shown
CC in the specification. It has been supplied in an electronic format from
CC WIPO.
XX
XQ Sequence 491 AA;

Query Match 100.0%; Score 2676; DB 7; Length 491;
Best Local Similarity 100.0%; Pred. No. 9.1e-215;
Matches 491; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MKRMVSWSPHKLTKMKHLLLLCVLVKSGQVNDNEEGFFSARGHRPLDKKREAPSRL 60
Db 1 MKRMVSWSPHKLTKMKHLLLLCVLVKSGQVNDNEEGFFSARGHRPLDKKREAPSRL 60
QY 61 PAPPISGGYRARPAPAKAAATQKVERKAPDAGGCLHADPDGLVLCPTGCQLEALLQOE 120
Db 61 PAPPISGGYRARPAPAKAAATQKVERKAPDAGGCLHADPDGLVLCPTGCQLEALLQOE 120
QY 121 RPIRSVDELNNVAVSQTSSSSFOYMYLLKDLQKQKQVKNENNVNYSSELEKHQ 180
Db 121 RPIRSVDELNNVAVSQTSSSSFOYMYLLKDLQKQKQVKNENNVNYSSELEKHQ 180
QY 181 LYIDTVNSNIATNLRLVLSILENRSKIOLKLESVSAQMEYCRTPCTVSCNIPVVSKE 240
Db 181 LYIDTVNSNIATNLRLVLSILENRSKIOLKLESVSAQMEYCRTPCTVSCNIPVVSKE 240
QY 241 CEEIRKGETSEMYLIQFDSSVKFPRVYCDMNTENGWTVIQNRQDGSVDFGRKWDPK 300
Db 241 CEEIRKGETSEMYLIQFDSSVKFPRVYCDMNTENGWTVIQNRQDGSVDFGRKWDPK 300
QY 301 QGFGNVAITDGNKYCGLPGEYWLNDKISQLTRMGPTTELLIEMEDWKGDKVKAHYGGFT 360
Db 301 QGFGNVAITDGNKYCGLPGEYWLNDKISQLTRMGPTTELLIEMEDWKGDKVKAHYGGFT 360
QY 361 VQNEANKYQISVKNYKGTAGNALMDGASQIMGNRTWTTHNGMFFSTYDRDNDGWLTSDD 420
Db 361 VQNEANKYQISVKNYKGTAGNALMDGASQIMGNRTWTTHNGMFFSTYDRDNDGWLTSDD 420
QY 421 RQCKSKEDGGGWWYNNRCHAANPNRYWGQYTWDMAXHGTDDGVVMMNWKGSWYSMRKM 480
Db 421 RQCKSKEDGGGWWYNNRCHAANPNRYWGQYTWDMAXHGTDDGVVMMNWKGSWYSMRKM 480
QY 481 SMKIRPPFPQQ 491
Db 481 SMKIRPPFPQQ 491

RESULT 4
AAR82243
ID AAR82243 standard; protein; 491 AA.
AC AAR82243;
DT 26-OCT-1996 (first entry)
DE Human fibrinogen B-beta chain protein.
KW Human fibrinogen B-beta chain; transgenic animal milk; treatment;
KW sheep beta-lactoglobulin signal peptide fusion protein;
KW surgical adhesive.
OS Homo sapiens.
PN WO9523868-A1.
XX

QY 121 RPIRNSVDLNNVAVSOTSSSFQYMYLLKDLWQKQKQVNDNENNVNYSSELEKHQ 180
 DB 121 RPIRNSVDLNNVAVSOTSSSFQYMYLLKDLWQKQKQVNDNENNVNYSSELEKHQ 180
 QY 181 LYIDETVNSNIATNLRLVLSILENLSKIQLKESDVSQAQMEYCRTPCTVSCNIPVVSQKE 240
 DB 181 LYIDETVNSNIPTNLRLVLSILENLSKIQLKESDVSQAQMEYCRTPCTVSCNIPVVSQKE 240
 QY 241 CEELIRKGETSEMYLIQDSSVKPYRVVCDMNTENGWTVIQRQDGSVDFGRKWDPYK 300
 DB 241 CEELIRKGETSEMYLIQDSSVKPYRVVCDMNTENGWTVIQRQDGSVDFGRKWDPYK 300
 QY 301 QFGNVATNTDGNVYGLPGEYWLGNDKISQLTRMGPTPELLIEMEDWKGDVKVAHYGGFT 360
 DB 301 QFGNVATNTDGNVYGLPGEYWLGNDKISQLTRMGPTPELLIEMEDWKGDVKVAHYGGFT 360
 QY 361 VQNEANKYQISVKNYRGTTAGNALMDGASQLMGENRTMTIHNGMFFSTYDRDNDGWLTSDF 420
 DB 361 VQNEANKYQISVKNYRGTTAGNALMDGASQLMGENRTMTIHNGMFFSTYDRDNDGWLTSDF 420
 QY 421 RKQCKEDGGWYNRCHAANPNRGYVWGQYTWDMAKHGTDGVTVMWKNKGSWYSMRKM 480
 DB 421 RKQCKEDGGWYNRCHAANPNRGYVWGQYTWDMAKHGTDGVTVMWKNKGSWYSMRKM 480
 QY 481 SMKIRPFFPQQ 491
 DB 481 SMKIRPFFPQQ 491

RESULT 6
 AAM79475
 ID AAM79475 standard; protein; 495 AA.
 XX
 AC AAM79475;
 XX
 DT 06-NOV-2001 (first entry)
 DE Human protein SEQ ID NO 3121.
 XX
 KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
 KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
 KW tissue growth factor; immunomodulatory; cancer; leukaemia;
 KW nervous system disorder; arthritis; inflammation.
 OS Homo sapiens.
 XX
 PN WO200157190-A2.
 XX
 PD 09-AUG-2001.
 XX
 PF 05-FEB-2001; 2001WO-US0004098.
 XX
 PR 03-FEB-2000; 2000US-00496914.
 XX
 PR 27-APR-2000; 2000US-00560875.
 XX
 PR 20-JUN-2000; 2000US-00598075.
 XX
 PR 19-JUL-2000; 2000US-00620325.
 XX
 PR 01-SEP-2000; 2000US-00654936.
 XX
 PR 15-SEP-2000; 2000US-00663561.
 XX
 PR 20-OCT-2000; 2000US-00693325.
 XX
 PR 30-NOV-2000; 2000US-00728422.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y;
 PI Ma Y, Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;
 PI Xue AJ, Yang Y, Wejhrman T, Goodrich R;
 XX
 WPI; 2001-476283/51.
 DR N-PSDB; AAK52608.
 DR
 XX Nucleic acids encoding polypeptides with cytokine-like activities, useful
 PT in diagnosis and gene therapy.

XX
 PS
 XX
 CC Claim 20; Page 257-258; 6221pp; English.
 CC The invention relates to polynucleotides (AAK51456-AAK53435) and the
 CC encoded polypeptides (AAM78323-AAM80302) that exhibit activity elating to
 CC cytokine, cell proliferation or cell differentiation or which may induce
 CC production of other cytokines in other cell populations. The
 CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
 CC peptide therapy. The polypeptides have various cytokine-like activities,
 CC e.g. stem cell growth factor activity, haematopoiesis regulating
 CC activity, tissue growth factor activity, immunomodulatory activity and
 CC activin/inhibin activity and may be useful in the diagnosis and/or
 CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
 CC inflammation. Note: Records for SEQ ID NO 2110 (AAK52581), 2111
 CC (AAK52582) and 3666 (AAM80020) are omitted as the relevant pages from the
 CC sequence listing were missing at the time of publication
 XX
 SQ Sequence 495 AA;
 Query Match 99.1%; Score 2651; DB 4; Length 495;
 Best Local Similarity 99.4%; Pred. No. 1.1e-212;
 Matches 486; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 QY 3 RMVVSFHLKTMKHLILLLLCVLVKSGVNDNBERGFFSARGHRPLDKKREAPSLRPA 62
 DB 7 RPSVSFHLKTMKHLILLLLCVLVKSGVNDNBERGFFSARGHRPLDKKREAPSLRPA 66
 QY 63 PPIISGGGYRARPAAATQKKVERKAPDAGCLHADPDLGVLCTPGCOLQEARLP 122
 DB 67 PPIISGGGYRARPAAATQKKVERKAPDAGCLHADPDLGVLCTPGCOLQEARLP 126
 QY 123 IRNSVDELNNVAVSOTSSSFQYMYLLKDLWQKQKQVNDNENNVNYSSELEKHQ 182
 DB 127 IRNSVDELNNVAVSOTSSSFQYMYLLKDLWQKQKQVNDNENNVNYSSELEKHQ 186
 QY 183 IDETVNSNIATNLRLVLSILENLSKIQLKESDVSQAQMEYCRTPCTVSCNIPVVSQKE 242
 DB 187 IDETVNSNIPTNLRLVLSILENLSKIQLKESDVSQAQMEYCRTPCTVSCNIPVVSQKE 246
 QY 243 EIIRKGETSEMYLIQDSSVKPYRVVCDMNTENGWTVIQRQDGSVDFGRKWDPYK 302
 DB 247 EIIRKGETSEMYLIQDSSVKPYRVVCDMNTENGWTVIQRQDGSVDFGRKWDPYK 306
 QY 303 FGNVATNTDGNVYGLPGEYWLGNDKISQLTRMGPTPELLIEMEDWKGDVKVAHYGGFT 362
 DB 307 FGNVATNTDGNVYGLPGEYWLGNDKISQLTRMGPTPELLIEMEDWKGDVKVAHYGGFT 366
 QY 363 NEANKYQISVKNYRGTTAGNALMDGASQLMGENRTMTIHNGMFFSTYDRDNDGWLTSDF 422
 DB 367 NEANKYQISVKNYRGTTAGNALMDGASQLMGENRTMTIHNGMFFSTYDRDNDGWLTSDF 426
 QY 423 QCKEDGGWYNRCHAANPNRGYVWGQYTWDMAKHGTDGVTVMWKNKGSWYSMRKM 482
 DB 427 QCKEDGGWYNRCHAANPNRGYVWGQYTWDMAKHGTDGVTVMWKNKGSWYSMRKM 486
 QY 483 KIRPFFPQQ 491
 DB 487 KIRPFFPQQ 495

RESULT 7
 AAM79476
 ID AAM79476 standard; protein; 495 AA.
 XX
 AC AAM79476;
 XX
 DT 06-NOV-2001 (first entry)
 XX
 DE Human protein SEQ ID NO 3122.
 XX
 KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
 KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
 KW tissue growth factor; immunomodulatory; cancer; leukaemia;

nervous system disorder; arthritis; inflammation.

Homo sapiens.

WO200157190-A2.

09-AUG-2001.

05-FEB-2001; 2001WO-US004098.

03-FEB-2000; 2000US-00496914.

27-APR-2000; 2000US-00560875.

20-JUN-2000; 2000US-00598075.

19-JUL-2000; 2000US-00620325.

01-SEP-2000; 2000US-00654936.

15-SEP-2000; 2000US-00663561.

20-OCT-2000; 2000US-00693325.

30-NOV-2000; 2000US-00728422.

(HYSE-) HYSEQ INC.

Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y;
Ma Y, Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;
Xue AJ, Yang Y, Wejhrman T, Goodrich R;
WPI; 2001-476283/51.
N-PSDB; AAK52609.

Nucleic acids encoding polypeptides with cytokine-like activities, useful in diagnosis and gene therapy.

Claim 20; Page 258; 6221pp; English.

The invention relates to polynucleotides (AAK51456-AAK53435) and the encoded polypeptides (AAM78323-AAM80302) that exhibit activity elating to cytokine, cell proliferation or cell differentiation or which may induce production of other cytokines in other cell populations. The polynucleotides and polypeptides are useful in gene therapy, vaccines or peptide therapy. The polypeptides have various cytokine-like activities, e.g. stem cell growth factor activity, haematopoiesis regulating activity, tissue growth factor activity, immunomodulatory activity and activity/inhibit activity and may be useful in the diagnosis and/or treatment of cancer, leukaemia, nervous system disorders, arthritis and

Query Match 99.1%; Score 2651; DB 4; Length 495;
Best Local Similarity 99.4%; Pred. No. 1.1e-212;
Matches 486; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

3 RMVWSFHLKTKMKHLLLLCVFLVKSQGVNDNEEGFSARGHRPLDKKREAPSIRPA 62
7 RPVSWSFHLKTKMKHLLLLCVFLVKSQGVNDNEEGFSARGHRPLDKKREAPSIRPA 66

63 PPIISGGYRARPAAKAAATQKKVERKAPDAGGCLHADPDGLVLCPTGCQLQEALLQERP 122
67 PPIISGGYRARPAAKAAATQKKVERKAPDAGGCLHADPDGLVLCPTGCQLQEALLQERP 126

123 IRNSVDLNNVRAVSQTSSTSSQYMYLLKDLWQKQKQVKDNNVNVNYSSELEKHLY 182
127 IRNSVDLNNVRAVSQTSSTSSQYMYLLKDLWQKQKQVKDNNVNVNYSSELEKHLY 186

183 IDTVNSNATNLRLVLSIENLRSKIQLESVDSVQMEYCRTPCTVSCNIPVWSKECE 242
187 IDTVNSNATNLRLVLSIENLRSKIQLESVDSVQMEYCRTPCTVSCNIPVWSKECE 246

243 EIIKKGSETSEMYLIQPDSSVKPRVYCDMNTGCGTWTIQNRQDGSVDFGRKWDPKQG 302
247 EIIKKGSETSEMYLIQPDSSVKPRVYCDMNTGCGTWTIQNRQDGSVDFGRKWDPKQG 306

303 FGNVATNTDGKNYCGLPGEYMLGNDKISQLTRMGPTTELLIEMEDWKGDKVKAHGGFTVQ 362
307 FGNVATNTDGKNYCGLPGEYMLGNDKISQLTRMGPTTELLIEMEDWKGDKVKAHGGFTVQ 366

363 NEANKYQISVVKYRGCTAGNALMDGASQLMGENRTWTIHNGMFFSTYDRDNDGWLTSDDPK 422
367 NEANKYQISVVKYRGCTAGNALMDGASQLMGENRTWTIHNGMFFSTYDRDNDGWLTSDDPK 426

423 QCSKEDGGGWYNNRCHAAANPNRGYYWGGQYTWDMAKHGTDDGVVWNNWKGWSYMRKXSM 482
427 QCSKEDGGGWYNNRCHAAANPNRGYYWGGQYTWDMAKHGTDDGVVWNNWKGWSYMRKXSM 486

483 KIRPFPPOQ 491
487 KIRPFPPOQ 495

RESULT 8
AAM79477
ID AAM79477 standard; protein; 495 AA.
XX
AC AAM79477;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human protein SEQ ID NO 3123.
XX
KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukaemia;
KW nervous system disorder; arthritis; inflammation.
XX
OS Homo sapiens.
XX
PN WO200157190-A2.
XX
PD 09-AUG-2001.
XX
PF 05-FEB-2001; 2001WO-US004098.
XX
PR 03-FEB-2000; 2000US-00496914.
PR 27-APR-2000; 2000US-00560875.
PR 20-JUN-2000; 2000US-00598075.
PR 19-JUL-2000; 2000US-00620325.
PR 01-SEP-2000; 2000US-00654936.
PR 15-SEP-2000; 2000US-00663561.
PR 20-OCT-2000; 2000US-00693325.
PR 30-NOV-2000; 2000US-00728422.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y;
PI Ma Y, Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;
PI Xue AJ, Yang Y, Wejhrman T, Goodrich R;
XX
XX WPI; 2001-476283/51.
DR N-PSDB; AAK52610.
XX
XX Nucleic acids encoding polypeptides with cytokine-like activities, useful in diagnosis and gene therapy.
PT
PS Claim 20; Page 258; 6221pp; English.
XX
XX The invention relates to polynucleotides (AAK51456-AAK53435) and the encoded polypeptides (AAM78323-AAM80302) that exhibit activity elating to cytokine, cell proliferation or cell differentiation or which may induce production of other cytokines in other cell populations. The polynucleotides and polypeptides are useful in gene therapy, vaccines or peptide therapy. The polypeptides have various cytokine-like activities, e.g. stem cell growth factor activity, haematopoiesis regulating activity, tissue growth factor activity, immunomodulatory activity and activity/inhibit activity and may be useful in the diagnosis and/or treatment of cancer, leukaemia, nervous system disorders, arthritis and inflammation. Note: Records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666 (AAM80020) are omitted as the relevant pages from the sequence listing were missing at the time of publication
XX
XX Sequence 495 AA;
SQ

CC inflammation. Note: Records for SEQ ID NO 2110 (AAK52581), 2111
CC (AAK52582) and 3666 (AAM80020) are omitted as the relevant pages from the
CC sequence listing were missing at the time of publication
XX
SQ Sequence 495 AA;

Query Match 99.1%; Score 2651; DB 4; Length 495;
Best Local Similarity 99.4%; Pred. No. 1.1e-212;
Matches 486; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
QY 3 RMVSWFHLKTKWKHLLLLCVLVKSGVNDNEEGFFSARGHRPLDKKREAPSRLPA 62
Db 7 RPYWSFHLKTKWKHLLLLCVLVKSGVNDNEEGFFSARGHRPLDKKREAPSRLPA 66
QY 63 PPISSGGYRPAKAAATOKKVERKAPDAGGCLHADPDGLVLCPTGCOLQOALLQOERP 122
Db 67 PPISSGGYRPAKAAATOKKVERKAPDAGGCLHADPDGLVLCPTGCOLQOALLQOERP 126
QY 123 IRNSVDELNNNEAVSQTSSTSSSFQYMYLLKDLWKQKQVKNENNVNYSSELEKHLY 182
Db 127 IRNSVDELNNNEAVSQTSSTSSSFQYMYLLKDLWKQKQVKNENNVNYSSELEKHLY 186
QY 183 IDTVNSNATNLRVLSILENRSKIQLKESDVSAQMEYCRTPCTVSCNIPVVSKECE 242
Db 187 IDTVNSNATNLRVLSILENRSKIQLKESDVSAQMEYCRTPCTVSCNIPVVSKECE 246
QY 243 EIIRKGETSEMYLIQPDSSVKPYRVYCDMNTENGWTVIQRQDGSVDFGRKWDYKQG 302
Db 247 EIIRKGETSEMYLIQPDSSVKPYRVYCDMNTENGWTVIQRQDGSVDFGRKWDYKQG 306
QY 303 FGNVATNTDGKNYCGLPGEYWLGNDSISQLTRMGPTTELLIEMEDWKGDKVKAHYGGFTVQ 362
Db 307 FGNVATNTDGKNYCGLPGEYWLGNDSISQLTRMGPTTELLIEMEDWKGDKVKAHYGGFTVQ 366
QY 363 NEANKYQISVKNYRGTAGNALMDGASQLMGENRTMTIHNGMFFSTYDRNDGWLTSDDRK 422
Db 367 NEANKYQISVKNYRGTAGNALMDGASQLMGENRTMTIHNGMFFSTYDRNDGWLTSDDRK 426
QY 423 QCSKEDGGGWWYNRCHAANPNRGYYWGGQYTWDMAXHGTDDGVVWNNWKGSWYMKKSM 482
Db 427 QCSKEDGGGWWYNRCHAANPNRGYYWGGQYTWDMAXHGTDDGVVWNNWKGSWYMKKSM 486
QY 483 IRPFPFQQ 491
Db 487 IRPFPFQQ 495

RESULT 9
ADE76868
ID ADE76868 standard; protein; 488 AA.
XX
AC ADE76868;
XX
DT 29-JAN-2004 (first entry)
XX
DE Human protein expressed in a liver disorder #11.
XX
KW human; liver disorder; hyperlipidaemia; hypertension; type II diabetes;
KW tumour; liver; inflammatory disorder; immune response disorder;
KW high-throughput screening; differential gene expression; gene therapy.
XX
OS Homo sapiens.

XX US2003108871-A1.
XX
XX 12-JUN-2003.
XX
XX 30-JUL-2001; 2001US-00919039.
XX
XX 28-JUL-2000; 2000US-0222113P.
XX
XX (KASE/) KASER M R.

PI Kaser MR;
XX WPI; 2004-031227/03.
DR N-PSDB; ADE76867.
XX
PT Composition comprising several cDNAs that are differentially expressed in
PT treated human C3A liver cell cultures, useful for treating liver
PT disorders.
XX
XX Claim 1; SEQ ID NO 33; 41pp; English.

PS The invention relates to a composition comprising several cDNAs that are
XX differentially expressed in a liver disorder. The composition is useful
CC for treating liver disorder such as hyperlipidaemia, hypertension, type
CC II diabetes, tumours of the liver and disorders of the inflammatory and
CC immune response. The composition is useful for a high-throughput method
CC of screening several molecules or compounds to identify a ligand which
CC specifically binds a cDNA. A protein encoded by the cDNA is useful for a
CC high-throughput method for using a protein to screen several molecules or
CC compounds to identify at least one ligand which specifically binds the
CC protein which involves combining the protein encoded by the cDNA with
CC several of molecules or compounds under conditions to allow specific
CC binding, and detecting specific binding between the protein and a
CC molecule or compound therefore identifying a ligand which specifically
CC binds the protein. The composition is useful for detecting and
CC quantifying differential gene expression, can be used in gene therapy, to
CC formulate prognosis and to design a treatment regimen and to monitor the
CC efficacy of treatment. The present sequence represents the amino acid
CC sequence of a protein encoded by a cDNA differentially expressed in a
CC liver disorder.

XX Sequence 488 AA;

Query Match 99.0%; Score 2648; DB 8; Length 488;
Best Local Similarity 99.4%; Pred. No. 2e-212;
Matches 485; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 4 MYWSFHLKTKWKHLLLLCVLVKSGVNDNEEGFFSARGHRPLDKKREAPSRLPAP 63
Db 1 MYWSFHLKTKWKHLLLLCVLVKSGVNDNEEGFFSARGHRPLDKKREAPSRLPAP 60
QY 64 PPISSGGYRPAKAAATOKKVERKAPDAGGCLHADPDGLVLCPTGCOLQOALLQOERP 123
Db 61 PPISSGGYRPAKAAATOKKVERKAPDAGGCLHADPDGLVLCPTGCOLQOALLQOERP 120
QY 124 IRNSVDELNNNEAVSQTSSTSSSFQYMYLLKDLWKQKQVKNENNVNYSSELEKHLY 183
Db 121 IRNSVDELNNNEAVSQTSSTSSSFQYMYLLKDLWKQKQVKNENNVNYSSELEKHLY 180
QY 184 IDTVNSNATNLRVLSILENRSKIQLKESDVSAQMEYCRTPCTVSCNIPVVSKECE 243
Db 181 IDTVNSNATNLRVLSILENRSKIQLKESDVSAQMEYCRTPCTVSCNIPVVSKECE 240
QY 244 EIIRKGETSEMYLIQPDSSVKPYRVYCDMNTENGWTVIQRQDGSVDFGRKWDYKQG 303
Db 241 EIIRKGETSEMYLIQPDSSVKPYRVYCDMNTENGWTVIQRQDGSVDFGRKWDYKQG 300
QY 304 GNVATNTDGKNYCGLPGEYWLGNDSISQLTRMGPTTELLIEMEDWKGDKVKAHYGGFTVQ 363
Db 301 GNVATNTDGKNYCGLPGEYWLGNDSISQLTRMGPTTELLIEMEDWKGDKVKAHYGGFTVQ 360
QY 364 EANKYQISVKNYRGTAGNALMDGASQLMGENRTMTIHNGMFFSTYDRNDGWLTSDDRK 423
Db 361 EANKYQISVKNYRGTAGNALMDGASQLMGENRTMTIHNGMFFSTYDRNDGWLTSDDRK 420
QY 424 QCSKEDGGGWWYNRCHAANPNRGYYWGGQYTWDMAXHGTDDGVVWNNWKGSWYMKKSM 483
Db 421 QCSKEDGGGWWYNRCHAANPNRGYYWGGQYTWDMAXHGTDDGVVWNNWKGSWYMKKSM 480
QY 484 IRPFPFQQ 491
Db 481 IRPFPFQQ 488

```

RESULT 10
AAM78491
ID AAM78491 standard; protein; 539 AA.
XX
AC AAM78491;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human protein SEQ ID NO 1153.
XX
KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukaemia;
KW nervous system disorder; arthritis; inflammation.
XX
OS Homo sapiens.
XX
PN WO200157190-A2.
XX
PD 09-AUG-2001.
XX
PF 05-FEB-2001; 2001WO-US004098.
XX
PR 03-FEB-2000; 2000US-00496914.
PR 27-APR-2000; 2000US-00560875.
PR 20-JUN-2000; 2000US-00598075.
PR 19-JUL-2000; 2000US-00620325.
PR 01-SEP-2000; 2000US-00654936.
PR 15-SEP-2000; 2000US-00663561.
PR 20-OCT-2000; 2000US-00693325.
PR 30-NOV-2000; 2000US-00728422.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y;
PI Ma Y, Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;
PI Xue AJ, Yang Y, Wejhrman T, Goodrich R;
XX
DR WPI; 2001-476283/51.
DR N-PSDB; AAK51624.
XX
XX Nucleic acids encoding polypeptides with cytokine-like activities, useful
in diagnosis and gene therapy.
XX
PS Claim 20; Page 3388-3389; 6221pp; English.
XX
XX The invention relates to polynucleotides (AAK51456-AAK53435) and the
encoded polypeptides (AAM78323-AAM80302) that exhibit activity elating to
cytokine, cell proliferation or cell differentiation or which may induce
production of other cytokines in other cell populations. The
polynucleotides and polypeptides are useful in gene therapy, vaccines or
peptide therapy. The polypeptides have various cytokine-like activities,
e.g. stem cell growth factor activity, haematopoiesis regulating
activity, tissue growth factor activity, immunomodulatory activity and
activin/inhibin activity and may be useful in the diagnosis and/or
treatment of cancer, leukaemia, nervous system disorders, arthritis and
inflammation. Note: Records for SEQ ID NO 2110 (AAK52581), 2111
(AAK52582) and 3666 (AAM80020) are omitted as the relevant pages from the
sequence listing were missing at the time of publication
XX
SQ Sequence 539 AA;
Query Match 98.5%; Score 2637; DB 4; Length 539;
Best Local Similarity 90.9%; Pred. No. 1.9e-211;
Matches 490; Conservative 0; Mismatches 1; Indels 48; Gaps 1;
QY 1 MKRMVSWSFHKLTKMTHLLILLICVFLVKSGQVNDNEE-----38
DB 1 MKRMVSWSFHKLTKMTHLLILLICVFLVKSGQVNDNEEQYRITKTRSENLTNYKIKE 60
QY 39 -----GFSNARGHRPLDKREAPSLRAPPPTSGGGYR 72

```


DE Human protein SEQ ID NO 1155.
XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukaemia;
KW nervous system disorder; arthritis; inflammation.
OS Homo sapiens.
XX
PN WO200157190-A2.
XX
PD 09-AUG-2001.
XX
XX
XX 05-FEB-2001; 2001WO-US004098.
XX
PR 03-FEB-2000; 2000US-00496914.
PR 27-APR-2000; 2000US-00560875.
PR 20-JUN-2000; 2000US-00598075.
PR 19-JUL-2000; 2000US-00620325.
PR 01-SEP-2000; 2000US-00654936.
PR 15-SEP-2000; 2000US-00663561.
PR 20-OCT-2000; 2000US-00693325.
PR 30-NOV-2000; 2000US-00728422.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y;
PI Ma Y, Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;
PI Xue AJ, Yang Y, Wejhrman T, Goodrich R;
XX
DR WPI; 2001-476283/51.
DR N-PSDB; AAK51626.
XX
PT Nucleic acids encoding polypeptides with cytokine-like activities, useful
PT in diagnosis and gene therapy.
XX
PS Claim 20; Page 3390-3391; 6221pp; English.
XX
CC The invention relates to polynucleotides (AAK51456-AAK53435) and the
CC encoded polypeptides (AAM78323-AAM80302) that exhibit activity elating to
CC cytokine, cell proliferation or cell differentiation or which may induce
CC production of other cytokines in other cell populations. The
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC peptide therapy. The polypeptides have various cytokine-like activities,
CC e.g. stem cell growth factor activity, haematopoiesis regulating
CC activity, tissue growth factor activity, immunomodulatory activity and
CC activin/inhibin activity and may be useful in the diagnosis and/or
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC inflammation. Note: Records for SEQ ID NO 2110 (AAK52581), 2111
CC (AAK52582) and 3666 (AAM80020) are omitted as the relevant pages from the
CC sequence listing were missing at the time of publication
XX
SQ Sequence 453 AA;
Query Match 31.0%; Score 2436; DB 4; Length 453;
Best Local Similarity 92.1%; Pred. No. 9.9e-195;
Matches 452; Conservative 0; Mismatches 1; Indels 38; Gaps 1;
QY 1 MKRMVSWSHKLTMTKHLILLLCVFLVKSQGVNDNEEGFFSARGHRPLDKKREAPSLR 60
DB 1 MKRMVSWSHKLTMTKHLILLLCVFLVKSQGVNDNEEGFFSARGHRPLDKKREAPSLR 60
QY 61 PAPPTSGGGYRARPAAKAAATKKVERKAPDAGGCLHADPDLGLVPTGQQLQEQEALLQOE 120
DB 61 PAPPTSGGGYRARPAAKAAATKKVERKAPDAGGCLHADPDLGLVPTGQQLQEQEALLQOE 120
QY 121 RPTNSVDELNNVVEAVSQTSSSFQMYLLKDLWKQKQKVDNNVNVNVEYSELEKHQ 180
DB 121 RPTNSVDELNNVVEAVSQTSSSFQMYLLKDLWKQKQKVDNNVNVNVEYSELEKHQ 180
QY 181 LYIDETVNSNIATNLRLVLSILENLSKIQKLESVDYSAQMEYCRTPCTVSCNIPVVSQKE 240
DB 181 LYIDETVNSNIPTNLRLVLSILENLSKIQKLESVDYSAQMEYCRTPCTVSCNIPVVSQK- 239

QY 241 CBEIIRKGGTSEMYLIQPDSSVKPYRYVCDMNTENGWTVIQNRQDGSVDGRKWDVPYK 300
DB 240 -----GWTVIQNRQDGSVDGRKWDVPYK 262
QY 301 QGFGNVATNTDGKNYCGLPGEYWLGNDKISQLTRMGPTTELLIEMEDWKDKYKAHYGGFT 360
DB 263 QGFGNVATNTDGKNYCGLPGEYWLGNDKISQLTRMGPTTELLIEMEDWKDKYKAHYGGFT 322
QY 361 VONEANKYQISVKNYRGTTAGNALMDGASQLMGENTMTIHNMGFTSTYDRDNDGWLTSDP 420
DB 323 VONEANKYQISVKNYRGTTAGNALMDGASQLMGENTMTIHNMGFTSTYDRDNDGWLTSDP 382
QY 421 RXQCSKEDGGGWYNNRCHAAANPNGRYWGGOYTTDMAKHGTDDGVVMMNWKGSWYSMRKM 480
DB 383 RXQCSKEDGGGWYNNRCHAAANPNGRYWGGOYTTDMAKHGTDDGVVMMNWKGSWYSMRKM 442
QY 481 SMKIRPFPFQQ 491
DB 443 SMKIRPFPFQQ 453
RESULT 14
ADI82107
ID ADI82107 standard; protein; 479 AA.
XX
AC ADI82107;
XX
DT 22-APR-2004 (first entry)
XX
DE Fibrinogen-beta-chain protein.
XX
KW kidney toxicity; T-kininogen; inter-alpha-inhibitor H4p heavy chain;
KW alpha-1-macroglobulin; apolipoprotein E; clusterin; complement C3;
KW complement C4; fibrinogen-alpha-chain; fibrinogen-beta-chain;
KW plasma retinol binding protein; renal tubular necrosis; drug development.
XX
OS Unidentified.
XX
PN WO2004005934-A2.
XX
PD 15-JAN-2004.
XX
PF 04-JUL-2003; 2003WO-GB002893.
XX
PR 04-JUL-2002; 2002GB-00015575.
PR 09-SEP-2002; 2002GB-00020879.
XX
PA (OXFO-) OXFORD GLYCOSCIENCES UK LTD.
XX
PI Ranasinghe-Bandara L;
XX
DR WPI; 2004-099412/10.
XX
PT Screening or diagnosing kidney toxicity, useful in monitoring
PT effectiveness of treatment for kidney toxicity or for screening and
PT developing of drugs, by detecting and quantifying T-kininogen and inter-
PT alpha-inhibitor H4p heavy chain.
XX
PS Disclosure; SEQ ID NO 10; 41pp; English.
XX
CC The invention comprises a method for screening or diagnosing kidney
CC toxicity. The method involves the use of the following proteins: T-
CC kininogen,inter-alpha-inhibitor H4p heavy chain, alpha-1-macroglobulin,
CC apolipoprotein E, clusterin, complement C3, complement C4, fibrinogen-
CC alpha-chain, fibrinogen-beta-chain, and plasma retinol binding protein.
CC The method of the invention is useful in screening or diagnosing kidney
CC toxicity (e.g. renal tubular necrosis). The method is also useful for
CC monitoring the effectiveness of treatment for kidney toxicity and for
CC screening and developing drugs. The present amino acid sequence
CC represents a protein that was used in the exemplification of the
CC invention.
XX

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 22, 2004, 15:08:20 ; Search time 2161.38 Seconds
(without alignments)
11849.778 Million cell updates/sec

Title: US-10-017-724-5_COPY_4000_8878

Perfect score: 4879

Sequence: 1 tcagaaatagtgtgatgag.....gattgcattgagccaagatc 4879

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4134886 seqs, 2624710521 residues

Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N_Geneseq_23Sep04:*

- 1: Geneseqn1980s:*
- 2: Geneseqn1990s:*
- 3: Geneseqn2000s:*
- 4: Geneseqn2001as:*
- 5: Geneseqn2001bs:*
- 6: Geneseqn2002as:*
- 7: Geneseqn2002bs:*
- 8: Geneseqn2003as:*
- 9: Geneseqn2003bs:*
- 10: Geneseqn2003cs:*
- 11: Geneseqn2003ds:*
- 12: Geneseqn2004as:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query %	Match	Length	ID	Description
1	4879	100.0	8878	2	AAT03852	Aat03852 Human fib
2	4879	100.0	8878	6	AAD45341	Aad45341 Human ant
3	4879	100.0	8878	9	AAD55117	Aad55117 Human FBG
4	4879	100.0	8878	11	ADP65553	Adp65553 Human fib
5	4879	100.0	8878	11	ADP65677	Adp65677 Human fib
6	4268.2	87.5	8674	8	ABX63814	Abx63814 Human cDN
7	2081.2	42.7	4043	6	ABX91986	Abx91986 Lung spec
8	971.8	19.9	4967	5	AAS60957	Aas60957 Human can
9	971.8	19.9	4967	5	ADL63650	Adl63650 Human ova
10	659.6	13.5	2553	4	AAS52608	Aas52608 Human pol
11	659.6	13.5	2553	4	AAS52609	Aas52609 Human pol
12	659.6	13.5	2553	4	AAK52610	Aak52610 Human pol
13	658.6	13.5	1862	4	AAK51626	Aak51626 Human pol
14	658.6	13.5	1976	4	AAK51625	Aak51625 Human pol
15	658.6	13.5	2120	4	AAK51624	Aak51624 Human pol
16	656.4	13.5	1918	11	ADP65102	Adp65102 Human fib
17	656.4	13.5	1918	11	ADP65031	Adp65031 Human fib
18	653.4	13.4	1883	11	ADP65676	Adp65676 Human fib
19	653.4	13.4	1883	11	ADP65552	Adp65552 Human fib
20	594	12.2	2553	4	AAK52608	Aak52608 Human pol
21	594	12.2	2553	4	AAK52609	Aak52609 Human pol

C	22	594	12.2	2553	4	AAK52610	Aak52610 Human pol
	23	454.6	9.3	455	3	AAC71409	Aac71409 Single nu
	24	454.6	9.3	455	3	AAC71418	Aac71418 Single nu
	25	454.6	9.3	455	3	AAC71443	Aac71443 Single nu
	26	447.6	9.2	448	3	AAC71412	Aac71412 Single nu
	27	447.6	9.2	448	3	AAC71421	Aac71421 Single nu
	28	430.6	8.8	431	3	AAC71427	Aac71427 Single nu
	29	397.2	8.1	414	3	ACH20725	Ach20725 Human adu
	30	390.4	8.0	404	9	ACH20470	Ach20470 Human adu
	31	387.8	7.9	415	9	ACH20471	Ach20471 Human adu
	32	387.8	7.9	1629	12	ADE76867	Ade76867 Human cDN
	33	369.4	7.6	399	9	ACH21083	Ach21083 Human adu
	34	368.4	7.6	382	9	ACH20960	Ach20960 Human adu
C	35	362.2	7.4	571	6	ABN96686	Abn96686 Gene #318
	36	338.2	6.9	375	6	ABN96703	Abn96703 Gene #320
	37	335.6	6.9	385	9	ACH21036	Ach21036 Human adu
	38	324.4	6.6	388	9	ACH21014	Ach21014 Human adu
	39	316.6	6.5	317	3	AAC71403	Aac71403 Single nu
	40	288.6	5.9	413	9	ACH20958	Ach20958 Human adu
	41	284.6	5.8	285	3	AAC71433	Aac71433 Single nu
	42	284.6	5.8	285	3	AAC71424	Aac71424 Single nu
	43	277.2	5.7	306	9	ACH20753	Ach20753 Human adu
	44	275.6	5.6	412	9	ACH20888	Ach20888 Human adu
C	45	264.4	5.4	357	6	ABN94253	Abn94253 Gene #751

ALIGNMENTS

RESULT 1

AAT03852

ID AAT03852 standard; DNA; 8878 BP.

XX AC AAT03852;

XX DT 26-OCT-1996 (first entry)

XX DE Human fibrinogen B-beta chain coding sequence.

XX KW Human fibrinogen B-beta chain; transgenic animal milk; treatment;

XX KW sheep beta-lactoglobulin signal peptide fusion protein;

XX KW surgical adhesive; ds.

XX OS Homo sapiens.

XX FH Key

FT misc_RNA

FT CDS

FT exon

FT intron

FT exon

FT intron

FT exon

FT intron

FT exon

FT intron

FT exon

FT intron

FT exon

FT intron

FT exon

FT intron

FT exon

FT intron

FT exon

FT intron

FT exon

FT intron

FT exon

FT intron

FT exon

FT intron

FT exon

FT intron

FT exon

FT intron

FT exon

FT intron

FT exon

FT intron

FT exon

FT intron

FT exon

```

FT      /*tag= i
FT      /note= "4"
FT      exon      5831..5944
FT      /*tag= j
FT      /note= "5"
FT      intron     5945..6632
FT      /*tag= k
FT      /note= "5"
FT      exon      6633..6758
FT      /*tag= l
FT      /note= "6"
FT      intron     6759..6966
FT      /*tag= m
FT      /note= "6"
FT      exon      6967..7252
FT      /*tag= n
FT      /note= "7"
FT      intron     7253..7870
FT      /*tag= o
FT      /note= "7"
FT      exon      7871..8102
FT      /*tag= p
FT      /note= "8"
FT      3'UTR      8103..8537
FT      /*tag= q
FT      misc_RNA   8538..8878
FT      /*tag= r
XX
PN WO9523868-A1.
XX
XX      08-SEP-1995.
XX
XX      01-MAR-1995; 95WO-US002648.
XX
XX      03-MAR-1994; 94US-00206176.
XX
XX      (ZYMO ) ZYMOGENETICS INC.
XX      (PHAR-) PHARM PROTEINS LTD.
XX
XX      Garner I, Dalrymple MA, Prunkard DE, Foster DC;
XX      WPI; 1995-320582/41.
XX      P-PSDB; AAR82243.
XX
XX      Production of fibrinogen in transgenic mammals - by introducing DNA
XX      segments into the germ line of a non-human mammal and collecting milk
XX      from female progeny.
XX
XX      Claim 9; Page 47-55; 99pp; English.
XX
XX      This sequence encodes the human fibrinogen B-beta chain. Preferably
XX      nucleotides 470-8100, and more preferably nucleotides 512-8100, are
XX      operably linked to the signal peptide, preferably of the sheep beta-
XX      lactoglobulin gene (see AAR03855) and, together with the fibrinogen A-
XX      alpha chain sequence (see AAR03853) and the gamma chain sequence (see
XX      AAR03854), is introduced into a fertilised egg or the germ line of a non-
XX      human animal, preferably a sheep, pig, goat or cattle at a molar ratio of
XX      0.5-1.0-0.5-1.0-0.5-1.0, respectively, for fusion protein gene expression
XX      in transgenic animal milk. The recombinantly produced active fibrinogen
XX      is useful in human and veterinary medicine, e.g. in the formulation of
XX      surgical adhesives, which also consist of Factor-XIII, and as a coating
XX      surface for polymeric articles, such as synthetic vascular grafts
XX
XX      Sequence 8878 BP; 2912 A; 1477 C; 1664 G; 2825 T; 0 U; 0 Other;
XX
XX      Query Match      100.0%; Score 4879; DB 2; Length 8878;
XX      Best Local Similarity 100.0%; Pred. No. 0;
XX      Matches 4879; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX      1 TCAGAAATAGTGTGATGAGTTAAATACAAATGTGGAAGCTGTTCCACAGACTCTCTCTT
XX      4000 TCAGAAATAGTGTGATGAGTTAAATACAAATGTGGAAGCTGTTCCACAGACTCTCTCTT
XX      Db

```

Db 5140 TGTGCTTCATCTCGAAACCTGAGAGGCAAAATACAAAAGTTAGAACTCTGATG 5199
Qy 1201 CTCAGCTCAAAATGGAATATGTCGACCCCATGCACTGCTGCAATATTCCTGTGGT 1260
Db 5200 CTCAGCTCAAAATGGAATATGTCGACCCCATGCACTGCTGCAATATTCCTGTGGT 5259
Qy 1261 GTCTGGCAAGGTAACTGATTCATAAACAATATTTAGAGAGTTCAGAGAACTCACAC 1320
Db 5260 GTCTGGCAAGGTAACTGATTCATAAACAATATTTAGAGAGTTCAGAGAACTCACAC 5319
Qy 1321 ACCAAAATAAGAGAAACAAACAACAACAAATGCTAAGTGGATTTTCCCAACAGATC 1380
Db 5320 ACCAAAATAAGAGAAACAAACAACAACAAATGCTAAGTGGATTTTCCCAACAGATC 5379
Qy 1381 ATAATGACATTTACAGTACATCATAAAAATATCTTTAGCCAGTGTGTTTTGGACTGGCCT 1440
Db 5380 ATAATGACATTTACAGTACATCATAAAAATATCTTTAGCCAGTGTGTTTTGGACTGGCCT 5439
Qy 1441 GGTGCAATTTGCTGTTTTGATGAGCAGGATGGGACAGGTAGTCCGAGGGTGGCTGAT 1500
Db 5440 GGTGCAATTTGCTGTTTTGATGAGCAGGATGGGACAGGTAGTCCGAGGGTGGCTGAT 5499
Qy 1501 GTGTGCATCTGCTGACTGGCTTGAACAGATGGCAGAACCAACAGATAGATGTAGAAAGTTTC 1560
Db 5500 GTGTGCATCTGCTGACTGGCTTGAACAGATGGCAGAACCAACAGATAGATGTAGAAAGTTTC 5559
Qy 1561 TCCATTTTGTGTGTTCTGGAGCTCATGGATATCCAGAGACAAAGGTGGAGAGAGC 1620
Db 5560 TCCATTTTGTGTGTTCTGGAGCTCATGGATATCCAGAGACAAAGGTGGAGAGAGC 5619
Qy 1621 TTTGTTTCATCTTACGAGATAAAGCTCCTCAAACTGGGTGGACTTACTTAAGATAAA 1680
Db 5620 TTTGTTTCATCTTACGAGATAAAGCTCCTCAAACTGGGTGGACTTACTTAAGATAAA 5679
Qy 1681 ATGAAAATCTAAATATTTGTTATATTTTCAAAGTCTATAATAACACACTCCTTAGTA 1740
Db 5680 ATGAAAATCTAAATATTTGTTATATTTTCAAAGTCTATAATAACACACTCCTTAGTA 5739
Qy 1741 ACTTATGTAATGTTATTTTAAAGAAATGGTGACTAAATAACAAGTAAATATGTCTATAAC 1800
Db 5740 ACTTATGTAATGTTATTTTAAAGAAATGGTGACTAAATAACAAGTAAATATGTCTATAAC 5799
Qy 1801 CCTGAAACATAATGTTGCTTACATTTGCAGATGTGAGGAAATATCAGGAAAGAGGT 1860
Db 5800 CCTGAAACATAATGTTGCTTACATTTGCAGATGTGAGGAAATATCAGGAAAGAGGT 5859
Qy 1861 GAAACATCTGAAATGATCTCATTCACCTGCAGCTTCTGTCAAACCGTATAGAGTATAC 1920
Db 5860 GAAACATCTGAAATGATCTCATTCACCTGCAGCTTCTGTCAAACCGTATAGAGTATAC 5919
Qy 1921 TGTGACATGAATACAGAAATGAGGTAAAGCTTTGCAGAGTTGTTGCACCTGTTGATCTGT 1980
Db 5920 TGTGACATGAATACAGAAATGAGGTAAAGCTTTGCAGAGTTGTTGCACCTGTTGATCTGT 5979
Qy 1981 AATTTATTTGATACCGTAAATGCGAGGAAACAGGCCAGGTGTGCTGCTCATACCTGT 2040
Db 5980 AATTTATTTGATACCGTAAATGCGAGGAAACAGGCCAGGTGTGCTGCTCATACCTGT 6039
Qy 2041 AATTTCCAGCACCTTTGGAGGCCAAAGTGGCTCATAGCTTTGAGCCTTAGGAGTTTGAACCT 2100
Db 6040 AATTTCCAGCACCTTTGGAGGCCAAAGTGGCTCATAGCTTTGAGCCTTAGGAGTTTGAACCT 6099
Qy 2101 AGCTGGGCAACATAATGAGACCTTAACCTCAACAAAAAATAACAAAAAATAACAAAAA 2160
Db 6100 AGCTGGGCAACATAATGAGACCTTAACCTCAACAAAAAATAACAAAAAATAACAAAAA 6159
Qy 2161 AAAAAATACGCTGTGTTGATGATGCTGTAGTCCAGCTATCCAGGAGGCTGAGA 2220
Db 6160 AAAAAATACGCTGTGTTGATGATGCTGTAGTCCAGCTATCCAGGAGGCTGAGA 6219
Qy 2221 TGGGAGATCACCTGAGGCCAACCTGGAGTCTTGATCATGCTACTGAACTGTAGCCTGG 2280

Db 6220 TGGGAGATCACCTGAGGCCCAACACCTGGAGTCTTGTGATCATGCTACTGAACCTGTAGCCTGG 6279
Qy 2281 GCAACAGAGTACTGAGATCCTGCTCAAAAABAAAATAATTTAAAAAGCCAGGAAAC 2340
Db 6280 GCAACAGAGTACTGAGATCCTGCTCAAAAABAAAATAATTTAAAAAGCCAGGAAAC 6339
Qy 2341 AAGACTTAGCTCTAACATCTAACATAGCTGACAAAGGAGTAAATTTGATGTGGAATTCAC 2400
Db 6340 AAGACTTAGCTCTAACATCTAACATAGCTGACAAAGGAGTAAATTTGATGTGGAATTCAC 6399
Qy 2401 CTGATATTTAAAAGTTATAAATATCTAATAATCTAATAATTTGGGTAAGATAAAGCATT 2460
Db 6400 CTGATATTTAAAAGTTATAAATATCTAATAATCTAATAATTTGGGTAAGATAAAGCATT 6459
Qy 2461 GCAGTTTCCAAGATTTTACAAGTTTACCTCTCATATTTATTTCCCTTATTTGCTCTATT 2520
Db 6460 GCAGTTTCCAAGATTTTACAAGTTTACCTCTCATATTTATTTCCCTTATTTGCTCTATT 6519
Qy 2521 TAGAGCAACAAATATATATACTAAATGGAATGGACAGGGATTCAGATATTTTCCAAAGT 2580
Db 6520 TAGAGCAACAAATATATATACTAAATGGAATGGACAGGGATTCAGATATTTTCCAAAGT 6579
Qy 2581 GACATATTTGCTGTGTTAATATATATGCTCTTTTTGTTTCTGTCAACAAAGGATGGAC 2640
Db 6580 GACATATTTGCTGTGTTAATATATATGCTCTTTTTGTTTCTGTCAACAAAGGATGGAC 6639
Qy 2641 AGTGATTCAGAACCGTCAAGCGGTAGTTGACTTTTGGCAGGAATGGATCCATATAA 2700
Db 6640 AGTGATTCAGAACCGTCAAGCGGTAGTTGACTTTTGGCAGGAATGGATCCATATAA 6699
Qy 2701 ACAGGGATTTGGAATGTTTGCACCAACACACAGATGGGAAGATTTACTGTGGCTACCAGG 2760
Db 6700 ACAGGGATTTGGAATGTTTGCACCAACACACAGATGGGAAGATTTACTGTGGCTACCAGG 6759
Qy 2761 TAAACGAAAGGATGCAAAAATAAAATCAATCTTATTTGAAATGGGATTTTTTTTAAATAAA 6819
Db 2821 AAACATTCATTTGTTGGAGCCTGTTTGGAGGTTTAAAGGAGTTTCCCTGACAAAATCT 2880
Qy 6820 AAACATTCATTTGTTGGAGCCTGTTTGGAGGTTTAAAGGAGTTTCCCTGACAAAATCT 6879
Db 2881 GGAAGCTAAAGATAAGGGAAGAAAGCCAGTTTATAGTTTCCCAAAATTTTTTTTGGTG 2940
Qy 6880 GGAAGCTAAAGATAAGGGAAGAAAGCCAGTTTATAGTTTCCCAAAATTTTTTTTGGTG 6939
Qy 2941 AGAGATTTTATTTGTTTCTTTTGGTGAATAATGCGCTTGGAAATGATAAATTAGCC 3000
Db 6940 AGAGATTTTATTTGTTTCTTTTAGGTGAATAATGCGCTTGGAAATGATAAATTAGCC 6999
Qy 3001 AGCTTACAGGATGGAGCCCAACAGACTTTTGCATAGAAATGGAGCTCGAAAGGAGACA 3060
Db 7000 AGCTTACAGGATGGAGCCCAACAGACTTTTGCATAGAAATGGAGCTCGAAAGGAGACA 7059
Qy 3061 AAGTAAAGGCTCACTATGAGGATTCACCTGTACAGAAATGAAGCCAAACAAATCCAGATCT 3120
Db 7060 AAGTAAAGGCTCACTATGAGGATTCACCTGTACAGAAATGAAGCCAAACAAATCCAGATCT 7119
Qy 3121 CAGTGAACAAATACAGAGAAACAGCGGTAAATGCGCTTCTATGATGGAGCATCTCAGCTGA 3180
Db 7120 CAGTGAACAAATACAGAGAAACAGCGGTAAATGCGCTTCTATGATGGAGCATCTCAGCTGA 7179
Qy 3181 TGGGAGAAAACAGGACCATGACCATTCACAAAGGATGTTCTTCCAGCATGATGACAGAG 3240
Db 7180 TGGGAGAAAACAGGACCATGACCATTCACAAAGGATGTTCTTCCAGCATGATGACAGAG 7239
Qy 3241 ACAATGACGCTGGTATGTTGGCTCTTCTGCTCTCTCTTAAAAATCACATAATATC 3300
Db 7240 ACAATGACGCTGGTATGTTGGCTCTTCTGCTCTCTCTTAAAAATCACATAATATC 7299
Qy 3301 ATTACTCAGAAATCATTAACAATAATTTTAAATAGTACCACTTCTCGGCACTTACTGTCA 3360
Db 7300 ATTACTCAGAAATCATTAACAATAATTTTAAATAGTACCACTTCTCGGCACTTACTGTCA 7359

QY	3361	GCCACTGTCCTAAAGCTCTTTATGATC	CACCTCGA	AAGCATTTCA	ACTATPAGG	TGACATT	3420
Db	7360	GCCACTGTCCTTAAGCTCTTTATGATC	CACCTCGA	AAGCATTTCA	ACTATPAGG	TGACATT	7419
QY	3421	CTTATTTCTCATTTTACAGATCAGATTT	TAGAGAGATTAC	GTGATTGTC	CAATGTCC	ACACACA	3480
Db	7420	CTTATTTCTCATTTTACAGATCAGATTT	TAGAGAGATTAC	GTGATTGTC	CAATGTCC	ACACACA	7479
QY	3481	ACTACCCAGAGATAAAACTAGAAATTT	GAGCAGTTACT	TTTCTG	AAATAAGCAGATT	TAG	3540
Db	7480	ACTACCCAGAGATAAAACTAGAAATTT	GAGCAGTTACT	TTTCTG	AAATAAGCAGATT	TAG	7539
QY	3541	ATAAATACCTATATCTCTATATTTCT	AAAGTGTGTG	AAAAC	TTCATTTTC	TCCAG	3600
Db	7540	ATAAATACCTATATCTCTATATTTCT	AAAGTGTGTG	AAAAC	TTCATTTTC	TCCAG	7599
QY	3601	GGTTCTCTGTACTATAAGGGTTGTTAA	AGCTATTATT	CCAGTATA	AAAGTAA	CAAAACACAGT	3660
Db	7600	GGTTCTCTGTACTATAAGGGTTGTTAA	AGCTATTATT	CCAGTATA	AAAGTAA	CAAAACACAGT	7659
QY	3661	CCCTAGATGGATTGCCACAAGGCCCG	AGTTATCTCT	TTTCTG	CTATAGGG	CACAGGAG	3720
Db	7660	CCCTAGATGGATTGCCACAAGGCCCG	AGTTATCTCT	TTTCTG	CTATAGGG	CACAGGAG	7719
QY	3721	GTCTTTGCGTATTAGTGTGACTCTAT	GTATAGCACC	CAAGGAAG	ACACTACT	TGTGCACA	3780
Db	7720	GTCTTTGCGTATTAGTGTGACTCTAT	GTATAGCACC	CAAGGAAG	ACACTACT	TGTGCACA	7779
QY	3781	CGAGTGTAGCAGTCTTTTATGGGTAA	TCTGTCA	AAACGTTAA	CTTGAC	CAACCGT	3840
Db	7780	CGAGTGTAGCAGTCTTTTATGGGTAA	TCTGTCA	AAACGTTAA	CTTGAC	CAACCGT	7839
QY	3841	TTCTTAATAACGCCAAAACATATTTT	CTTC	TTCAGGTTAA	CAATCAGAT	CCCGAAAAACAGT	3900
Db	7840	TTCTTAATAACGCCAAAACATATTTT	CTTC	TTCAGGTTAA	CAATCAGAT	CCCGAAAAACAGT	7899
QY	3901	CTAAGAAGACGGTGGTGGATATTAAT	AGATGTCTAT	GACGCC	CAATCC	CAAAACGGCA	3960
Db	7900	CTAAGAAGACGGTGGTGGATATTAAT	AGATGTCTAT	GACGCC	CAATCC	CAAAACGGCA	7959
QY	3961	GATACTACTGGGGTGGACAGTACAC	CTCGGCACAT	TGGCA	AGCATGG	CACAGATG	4020
Db	7960	GATACTACTGGGGTGGACAGTACAC	CTCGGCACAT	TGGCA	AGCATGG	CACAGATG	8019
QY	4021	TAGTATGGATGAATTTGGAAGGGG	TCATGTACT	CAATPAGG	AAAGATG	AGTATGA	4080
Db	8020	TAGTATGGATGAATTTGGAAGGGG	TCATGTACT	CAATPAGG	AAAGATG	AGTATGA	8079
QY	4081	GGCCCTTCTTCCACACAGATAGTCC	CCCATAGT	AGATT	TTTGGCT	CTTCTGTAT	4140
Db	8080	GGCCCTTCTTCCACACAGATAGTCC	CCCATAGT	AGATT	TTTGGCT	CTTCTGTAT	8139
QY	4141	AACATTTTGTACATTAATGTTATT	TGGAATTTTCT	TTTCAT	ACATTA	TATTTCTCT	4200
Db	8140	AACATTTTGTACATTAATGTTATT	TGGAATTTTCT	TTTCAT	ACATTA	TATTTCTCT	8199
QY	4201	CTCAGCAGAGGTGAGTGTGACTTTT	TGAAAAAG	TATAGG	ATAATTA	CAATTA	4260
Db	8200	CTCAGCAGAGGTGAGTGTGACTTTT	TGAAAAAG	TATAGG	ATAATTA	CAATTA	8259
QY	4261	CACATGATTTTCTTTTGTCTTTCT	TCTATTTCT	TGTCT	CACCA	AGAGTAA	4320
Db	8260	CACATGATTTTCTTTTGTCTTTCT	TCTATTTCT	TGTCT	CACCA	AGAGTAA	8319
QY	4321	AGTTTTGACAGAGTTGGTGTTCAT	AAATTTTC	AGTTT	TGAT	TGCGA	4380
Db	8320	AGTTTTGACAGAGTTGGTGTTCAT	AAATTTTC	AGTTT	TGAT	TGCGA	8379
QY	4381	AAGGAAGAGGGGCTTTTATCTTGT	CTGTAGG	AAAA	CCCATG	ACGGA	4440
Db	8380	AAGGAAGAGGGGCTTTTATCTTGT	CTGTAGG	AAAA	CCCATG	ACGGA	8439

QY	4441	GTTTAAAGATCCACTTTTAAACATATATTTATTTATGTAGATCTGTCAAAGAAAACCTTC	4500
Db	8440	GTTTAAAGATCCACTTTTAAACATATATTTATTTATGTAGATCTGTCAAAGAAAACCTTC	8499
QY	4501	CAAAAAGATTATTAATTTAAACACAGACTCTGTTGCAATAAGTTAAATGTTTCTTGTTTTG	4560
Db	8500	CAAAAAGATTATTAATTTAAACACAGACTCTGTTGCAATAAGTTAAATGTTTCTTGTTTTG	8559
QY	4561	TAATCCACACATTCATGATTTAGGCTTTTGCACTTTGTAAGGAAGGAGAACGTTCAAC	4620
Db	8560	TAATCCACACATTCATGATTTAGGCTTTTGCACTTTGTAAGGAAGGAGAACGTTCAAC	8619
QY	4621	CTCAAAATAGCTAATAAACCGGTCTTGTAATATTTGAAGATTTAAATCTGACTTAGGACG	4680
Db	8620	CTCAAAATAGCTAATAAACCGGTCTTGTAATATTTGAAGATTTAAATCTGACTTAGGACG	8679
QY	4681	GGCACGGTGGCTCACGACTATAATCCCAACACTTTTGGAGGCTGAGCGCGGCGTCAAA	4740
Db	8680	GGCACGGTGGCTCACGACTATAATCCCAACACTTTTGGAGGCTGAGCGCGGCGTCAAA	8739
QY	4741	GGTCAGAGTTCAAGACCGCTTGACCAATATGTTGAACCCCATCTCTACTAAAAATAC	4800
Db	8740	GGTCAGAGTTCAAGACCGCTTGACCAATATGTTGAACCCCATCTCTACTAAAAATAC	8799
QY	4801	AAAAATTAGCCAGCGCTGGTGGCAGGTGCTGTAGTCCCAGCTAGCCTGTGAGGTGGAG	4860
Db	8800	AAAAATTAGCCAGCGCTGGTGGCAGGTGCTGTAGTCCCAGCTAGCCTGTGAGGTGGAG	8859
QY	4861	ATTGCATTGAGCAAGATC	4879
Db	8860	ATTGCATTGAGCAAGATC	8878
RESULT 2			
AAD45341			
ID	AAD45341 standard; DNA; 8878 BP.		
XX	AAD45341;		
AC	27-DEC-2002 (first entry)		
DT	Human anti-microbial protein encoding DNA #4.		
DE	Human; feed improvement; plant-derived feed; antibiotic; additive;		
XX	anti-microbial protein; poultry; alpha-1-antitrypsin; flour; malt; ds.		
KW	Homo sapiens.		
OS	WO200263975-A2.		
FN	22-AUG-2002.		
XX	14-FEB-2002; 2002WO-US0004919.		
PF	14-FEB-2001; 2001US-0269188P.		
XX	02-MAY-2001; 2001US-00847232.		
PR	(VENT-) VENTRIA BIOSCIENCE.		
XX	Huang N, Rodriguez RL, Hagie FE;		
PI	WFI; 2002-682708/73.		
XX	Improved feed for production animals, comprising plant-derived feed		
PT	ingredients, and seed composition containing flour, extract, or malt from		
PT	mature monocot seeds and heterologous seed-produced anti-microbial		
PT	proteins.		
XX	Claim 21; Page 167-170; 175pp; English.		
PS	The invention relates to improved feed for production animals, comprising		
XX	one or more plant-derived feed ingredients, substantially unsupplemented		
CC	with small-molecule antibiotics and as an additive a seed composition		
CC	comprising plant-derived feed ingredients, and seed composition containing		
CC	flour, extract, or malt from mature monocot seeds and heterologous seed-produced anti-microbial		
CC	proteins.		
CC	Claim 21; Page 167-170; 175pp; English.		

CC containing a flour, extract or malt obtained from mature monocot seeds
CC and one or more heterologous seed-produced anti-microbial proteins in
CC substantially unpurified form. The invention is useful as a feed for
CC production animals such as poultry and hoofed farm animals. The present
CC sequence is human anti-microbial protein encoding DNA. This sequence is
XX used in the invention
SQ

Sequence 8878 BP; 2912 A; 1478 C; 1663 G; 2825 T; 0 U; 0 Other;

Query Match 100.0%; Score 4879; DB 6; Length 8878;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 4879; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCAGAAATAGTGTGATGAGTTAAATAACAAGTGGAGCTGTTCCAGAGCTCTCTTT 60
DB 4000 TCAGAAATAGTGTGATGAGTTAAATAACAAGTGGAGCTGTTCCAGAGCTCTCTTT 4059

QY 61 CTTCCCTTCAGTACATGATTTTGGCTGAAGACCTGTGGCAAGAGGCGAAGCAAGTAA 120
DB 4060 CTTCCCTTCAGTACATGATTTTGGCTGAAGACCTGTGGCAAGAGGCGAAGCAAGTAA 4119

QY 121 AAGGTAGATATCTGTGCTTCCATTCGATTTTCAGCTTATAAATTTGAAACCGTTAGAC 180
DB 4120 AAGGTAGATATCTGTGCTTCCATTCGATTTTCAGCTTATAAATTTGAAACCGTTAGAC 4179

QY 181 TGCCACGGAAGTGCATGGTGTGGAAGATTAACATTTCTGGGTAGTGAATAGCATTTCA 240
DB 4180 TGCCACGGAAGTGCATGGTGTGGAAGATTAACATTTCTGGGTAGTGAATAGCATTTCA 4239

QY 241 TACGGCTTTTGGGACCTTCCCTCGCAACTTGGCAGATAGCACTATTTCAGCTCTTATTC 300
DB 4240 TACGGCTTTTGGGACCTTCCCTCGCAACTTGGCAGATAGCACTATTTCAGCTCTTATTC 4299

QY 301 CAGTCTGACATCAGCAAGTGTGATTTTCTATGAAATTTCTACTAGCTCTCTTATTTTA 360
DB 4300 CAGTCTGACATCAGCAAGTGTGATTTTCTATGAAATTTCTACTAGCTCTCTTATTTTA 4359

QY 361 AGTATACAGAAACCTGTGACCTCAGAGATTAATTTACAGAGTGGAAAAAACCCTAG 420
DB 4360 AGTATACAGAAACCTGTGACCTCAGAGATTAATTTACAGAGTGGAAAAAACCCTAG 4419

QY 421 CATTTATAGTTTAAATTTGAGGTTTGAATGAGAGAGTTATCCATAATATATTTCAATT 480
DB 4420 CATTTATAGTTTAAATTTGAGGTTTGAATGAGAGAGTTATCCATAATATATTTCAATT 4479

QY 481 GTGTTGTGATTAATGACACTAACTGTGAATCTTGAGGTCAAGATTTGAGTGTGTTG 540
DB 4480 GTGTTGTGATTAATGACACTAACTGTGAATCTTGAGGTCAAGATTTGAGTGTGTTG 4539

QY 541 ACTTGGTGTGAGGAAACAGCTAGTGGTGGCTGGCAGGCACTCTCAGTGAGTAGCA 600
DB 4540 ACTTGGTGTGAGGAAACAGCTAGTGGTGGCTGGCAGGCACTCTCAGTGAGTAGCA 4599

QY 601 TACCACAGTTGGAAATTTTCAAGAAATCAAGGAATCATGACATCTTATAAATTTTCA 660
DB 4600 TACCACAGTTGGAAATTTTCAAGAAATCAAGGAATCATGACATCTTATAAATTTTCA 4659

QY 661 AGGTTCTGTCTACTTATGTAATGGAATAATAATCAAGCATATPCCAATCTGTAAGAT 720
DB 4660 AGGTTCTGTCTACTTATGTAATGGAATAATAATCAAGCATATPCCAATCTGTAAGAT 4719

QY 721 TGAACCTTCTCAGATGGAAGACCCCAATCTGTTCTCTCTCTTTCCTCCACCAAGAAA 780
DB 4720 TGAACCTTCTCAGATGGAAGACCCCAATCTGTTCTCTCTCTTTCCTCCACCAAGAAA 4779

QY 781 TAAACAACTTATTTCTATCTGACACAACTTTTAGCGGTATACCTATGTAATTTA 840
DB 4780 TAAACAACTTATTTCTATCTGACACAACTTTTAGCGGTATACCTATGTAATTTA 4839

QY 841 CTAGTATGTTGTTAGGATTTTATGTTAAATTTGTTATATGTCATGCGCCAAATCATTTCCAC 900
DB 4840 CTAGTATGTTGTTAGGATTTTATGTTAAATTTGTTATATGTCATGCGCCAAATCATTTCCAC 4899

QY 901 TAAATATGACATATATATCATAACTGCTTGGTGATAGCTCAGTGTTTAATAGTTTATTC 960
DB 4900 TAAATATGACATATATATCATAACTGCTTGGTGATAGCTCAGTGTTTAATAGTTTATTC 4959

QY 961 AGAAATCAAAATTTGATAGTTTAAATACATTTAGTTTATGAGCAAAATGCTAACTATT 1020
DB 4960 AGAAATCAAAATTTGATAGTTTAAATACATTTAGTTTATGAGCAAAATGCTAACTATT 5019

QY 1021 TCTACATTAATTTCAITTTTCCAGATAATGAATAATGATGACTCAATGACTCTCTCAGAACT 1080
DB 5020 TCTACATTAATTTCAITTTTCCAGATAATGAATAATGATGACTCAATGACTCTCTCAGAACT 5079

QY 1081 GGAAAGCACCAATATATATATAGATGAGACTGTGAATAGCAATATCCCAACTAACCTTCG 1140
DB 5080 GGAAAGCACCAATATATATATAGATGAGACTGTGAATAGCAATATCCCAACTAACCTTCG 5139

QY 1141 TGTGCTTTCGTTTCAATCTCGGAAACCTGAGAGCAAAATACAAAAGTTAGAAATCTGATGT 1200
DB 5140 TGTGCTTTCGTTTCAATCTCGGAAACCTGAGAGCAAAATACAAAAGTTAGAAATCTGATGT 5199

QY 1201 CTCAGCTCAAAATGGAATATTTGTCGACCCCATGACTGTGAGTTGCAATATTTCTGTGTTG 1260
DB 5200 CTCAGCTCAAAATGGAATATTTGTCGACCCCATGACTGTGAGTTGCAATATTTCTGTGTTG 5259

QY 1261 GTCTGCAAAAGTAACTGATTTCAATAAATATTTTGTAGAGGTTCCAGAGAACTCACAC 1320
DB 5260 GTCTGCAAAAGTAACTGATTTCAATAAATATTTTGTAGAGGTTCCAGAGAACTCACAC 5319

QY 1321 ACCAAAAATAGAGAAACCAACCAACAAAATGCTAAAGTGGATTTTCCCAACAGATTC 1380
DB 5320 ACCAAAAATAGAGAAACCAACCAACAAAATGCTAAAGTGGATTTTCCCAACAGATTC 5379

QY 1381 ATAATGACATTAACAGTACATCATATAAATAATCTTGTAGGCAAGTTGTTTGGACTGGCCT 1440
DB 5380 ATAATGACATTAACAGTACATCATATAAATAATCTTGTAGGCAAGTTGTTTGGACTGGCCT 5439

QY 1441 GGTGCAATTTGCTGTTTGTAGTGGAGGAGTGGGACAGTGGTGTAGTCCCAGGGTGGCTGAT 1500
DB 5440 GGTGCAATTTGCTGTTTGTAGTGGAGGAGTGGGACAGTGGTGTAGTCCCAGGGTGGCTGAT 5499

QY 1501 GTGTGCACTCTGCTTACTGGCTTGAACAGATGCGAGAACCCACAGATAGATGTAGAGTTTC 1560
DB 5500 GTGTGCACTCTGCTTACTGGCTTGAACAGATGCGAGAACCCACAGATAGATGTAGAGTTTC 5559

QY 1561 TCCATTTTGTGTGTTCTGGAGCTCATGGAATATCCAGAGACACAAAAGTGGAGAGAGC 1620
DB 5560 TCCATTTTGTGTGTTCTGGAGCTCATGGAATATCCAGAGACACAAAAGTGGAGAGAGC 5619

QY 1621 TTTGTTTCATCTCTTAGCAGATTAACGCTCCCAAACTGGGTTGGACTTACTTAAAGTAA 1680
DB 5620 TTTGTTTCATCTCTTAGCAGATTAACGCTCCCAAACTGGGTTGGACTTACTTAAAGTAA 5679

QY 1681 ATGAAATCTTAATATTTGTTTATTTTCAAAGGCTCTATAATAACACACTCTCTTAGTA 1740
DB 5680 ATGAAATCTTAATATTTGTTTATTTTCAAAGGCTCTATAATAACACACTCTCTTAGTA 5739

QY 1741 ACTTATGTAATGTTATTTTAAAGAAATTTGTTGACTAAATACAAAAGTAAATTTATGCTAAAC 1800
DB 5740 ACTTATGTAATGTTATTTTAAAGAAATTTGTTGACTAAATACAAAAGTAAATTTATGCTAAAC 5799

QY 1801 CCCTGAAATATGTTGTTCTTACATTTTCGAGAACTGGAGAAATTTATCAGGAAAGGAGGT 1860
DB 5800 CCCTGAAATATGTTGTTCTTACATTTTCGAGAACTGGAGAAATTTATCAGGAAAGGAGGT 5859

QY 1861 GAAACATCTGAAATCTATCTCAATCTCAACCTGACAGTTCTGTCAAACCGTATAGAGTATAC 1920
DB 5860 GAAACATCTGAAATCTATCTCAATCTCAACCTGACAGTTCTGTCAAACCGTATAGAGTATAC 5919

QY 1921 TGTGATGTAATGTAAGAAATGAGGTAAGCTTTTCGACAGTTGTTGAACTGTTGATCTGT 1980
DB 5920 TGTGATGTAATGTAAGAAATGAGGTAAGCTTTTCGACAGTTGTTGAACTGTTGATCTGT 5979

QY 1981 AATTATTTGGATACCGTAAATGCCAGGAACCAAGCCAGGTGTGTTGGCTCATACCTGT 2040

Db 5980 AATTATTTGGATACCGTAAATGCCAGGAAACAGGCCAGGTGTGGTGGCTCATACCTGT 6039
QY 2041 AATTCACGACCTTGGGAGGCCAAGTGGCTGATAGCTTGAGCCCTAGGAGTTTCAAAC 2100
Db 6040 AATTCACGACCTTGGGAGGCCAAGTGGCTGATAGCTTGAGCCCTAGGAGTTGAAACT 6099
QY 2101 AGCTGGGCAACATAATAGAGCCCTAACTCTACAAAATAAAAAAATAACCAAAAAAAA 2160
Db 6100 AGCTGGGCAACATAATAGAGCCCTAACTCTACAAAATAAAAAAATAACCAAAAAAAA 6159
QY 2161 AAAAAAATCAGCTGTGGTGTAGTATGTGCTGATGCTAGTCCAGCTATCCAGAGGCTGAGA 2220
Db 6160 AAAAAAATCAGCTGTGGTGTAGTATGTGCTGATGCTAGTCCAGCTATCCAGAGGCTGAGA 6219
QY 2221 TGGGAGATCAGCTGAGGCCCAACCTGGAGTCTTGATCATGCTAGTCCAGCTATCCAGAGGCTG 2280
Db 6220 TGGGAGATCAGCTGAGGCCCAACCTGGAGTCTTGATCATGCTAGTCCAGCTATCCAGAGGCTG 6279
QY 2281 GCAACAGAGGATAGTGAGATCCTGTCTCAAAAAAATAAATAAATAAAGCCAGGAAC 2340
Db 6280 GCAACAGAGGATAGTGAGATCCTGTCTCAAAAAAATAAATAAATAAAGCCAGGAAC 6339
QY 2341 AAGACTTAGCTCTAATCATCTAATAGTCTGACAAAGAGATTAATTTGATGTGGAATTCAC 2400
Db 6340 AAGACTTAGCTCTAATCATCTAATAGTCTGACAAAGAGATTAATTTGATGTGGAATTCAC 6399
QY 2401 CTGATATTTAAAGTTATATAATCTATATTTCAATTTGGGGTAAAGATAAAGCACTT 2460
Db 6400 CTGATATTTAAAGTTATATAATCTATATTTCAATTTGGGGTAAAGATAAAGCACTT 6459
QY 2461 GCAGTTTCCAAAGATTTCAAGTTTACCTCTCATATTTATTTCTTATTTGTCATTT 2520
Db 6460 GCAGTTTCCAAAGATTTCAAGTTTACCTCTCATATTTATTTCTTATTTGTCATTT 6519
QY 2521 TAGACACCAATATATATAATGGAATGGACAGGATTCAGATATTTATTTCAAAGT 2580
Db 6520 TAGACACCAATATATATAATGGAATGGACAGGATTCAGATATTTATTTCAAAGT 6579
QY 2581 GACATTTTGTGTGTGTTAATATATCTCTTTTGTCTCTGCAACCAAGAGTGGAC 2640
Db 6580 GACATTTTGTGTGTGTTAATATATCTCTTTTGTCTCTGCAACCAAGAGTGGAC 6639
QY 2641 AGTGATTCAGAACCGTCAAGCGGTAGTGTGACTTTGGCAGAAATGGATCCATATA 2700
Db 6640 AGTGATTCAGAACCGTCAAGCGGTAGTGTGACTTTGGCAGAAATGGATCCATATA 6699
QY 2701 ACAGGATTTGGAAATGTGCAACCAACAGATGGGAAGATTAATCTGTGGCTTACCAG 2760
Db 6700 ACAGGATTTGGAAATGTGCAACCAACAGATGGGAAGATTAATCTGTGGCTTACCAG 6759
QY 2761 TAAAGAACAGGATGCAAAATAAATCATTTCTATTTGAAATGGGATTTTTTTTAAATAAA 2820
Db 6760 TAAAGAACAGGATGCAAAATAAATCATTTCTATTTGAAATGGGATTTTTTTTAAATAAA 6819
QY 2821 AAACATTCATTTGGAAGCTGTTTTAGGAGTTTAAAGAGTTTCCGCAAAAAATGT 2880
Db 6820 AAACATTCATTTGGAAGCTGTTTTAGGAGTTTAAAGAGTTTCCGCAAAAAATGT 6879
QY 2881 GGAAGCTAAAGATAGGAGGAGGAGGAGGAGTTTATTTTCCCAAAATTTATTTTGGT 2940
Db 6880 GGAAGCTAAAGATAGGAGGAGGAGGAGGAGTTTATTTTCCCAAAATTTATTTTGGT 6939
QY 2941 AGAGATTTTATTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 3000
Db 6940 AGAGATTTTATTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 6999
QY 3001 AGCTTACAGGATGGGACCCACAGAACTTTTGAATAAAGTGGAGTCTGGAAGAGACA 3060
Db 7000 AGCTTACAGGATGGGACCCACAGAACTTTTGAATAAAGTGGAGTCTGGAAGAGACA 7059
QY 3061 AAGTAAAGCTCAGCTATGAGGATTCATCTGACAGATTCAGCCCAACCAATACAGATCT 3120

Db 7060 AAGTAAAGGCTCACTATGGAGGATTCATCTGTACAGAAATGAAGCCAAATAACAGATCT 7119
QY 3121 CAGTGAAACAAATACAGAGGAAACAGCCGGTAAATGCCCTCATGGATGGAGCATCTCAGCTGA 3180
Db 7120 CAGTGAAACAAATACAGAGGAAACAGCCGGTAAATGCCCTCATGGATGGAGCATCTCAGCTGA 7179
QY 3181 TGGGAGAAAAACAGGACCAATGACCAATTCACACGGCATGTTCTTACGACGCTATGACAGAG 3240
Db 7180 TGGGAGAAAAACAGGACCAATGACCAATTCACACGGCATGTTCTTACGACGCTATGACAGAG 7239
QY 3241 ACAATGACGGCTGTGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGG 3300
Db 7240 ACAATGACGGCTGTGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGG 7299
QY 3301 ATTACTCAGAAATCAATTAACAATTTTTTAATAGCTACCACTTCTCTGGGCACTTACTGTCA 3360
Db 7300 ATTACTCAGAAATCAATTAACAATTTTTTAATAGCTACCACTTCTCTGGGCACTTACTGTCA 7359
QY 3361 GCCACTGTCTTAAGCTCTTTATGCAATCACTCGAAGCAATTTCAACTATAAGGTAGACATTT 3420
Db 7360 GCCACTGTCTTAAGCTCTTTATGCAATCACTCGAAGCAATTTCAACTATAAGGTAGACATTT 7419
QY 3421 CTTATTTCTCAATTTACAGATGAGATTTAGAGAGATTAAGTGTGGTGTGGTGTGGTGTGGTGTGG 3480
Db 7420 CTTATTTCTCAATTTACAGATGAGATTTAGAGAGATTAAGTGTGGTGTGGTGTGGTGTGGTGTGG 7479
QY 3481 ACTACCCAGAGATAAAACTAGAAATTTGAGCACAGTTACTTTCTGAAATTAATGAGCATTTAG 3540
Db 7480 ACTACCCAGAGATAAAACTAGAAATTTGAGCACAGTTACTTTCTGAAATTAATGAGCATTTAG 7539
QY 3541 ATAAATACCTATCTCTATATTTCTAAAGTGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGG 3600
Db 7540 ATAAATACCTATCTCTATATTTCTAAAGTGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGG 7599
QY 3601 GGTCTCTGTACTAAAGGTTGTAAAGCTTATTTCCAGTATAAGTAAACAAACACAGT 3660
Db 7600 GGTCTCTGTACTAAAGGTTGTAAAGCTTATTTCCAGTATAAGTAAACAAACACAGT 7659
QY 3661 CCCTAGATGGATTCGCACAAAGGCCAGTTATCTCTTTCTTGTCTATAGGSCACAGGAG 3720
Db 7660 CCCTAGATGGATTCGCACAAAGGCCAGTTATCTCTTTCTTGTCTATAGGSCACAGGAG 7719
QY 3721 GTCTTTGTGTATTTAGTGTGACTCTATGTATAGCAACCAAGGAAAGACTACTGTGSCACA 3780
Db 7720 GTCTTTGTGTATTTAGTGTGACTCTATGTATAGCAACCAAGGAAAGACTACTGTGSCACA 7779
QY 3781 CGAGTGTAGCAGTCTTTTATGSGTAACTCTGCAAAACGTAACCTTGACACCCAGTGTCTGT 3840
Db 7780 CGAGTGTAGCAGTCTTTTATGSGTAACTCTGCAAAACGTAACCTTGACACCCAGTGTCTGT 7839
QY 3841 TTCTTAATAACGCCAAACACATTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 3900
Db 7840 TTCTTAATAACGCCAAACACATTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 7899
QY 3901 CTAAAGAAAGACGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGG 3960
Db 7900 CTAAAGAAAGACGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGG 7959
QY 3961 GATACCTAGTGG 4020
Db 7960 GATACCTAGTGG 8019
QY 4021 TAGTATGGAATTTGGAAGGGGTCAATGTTACTCAATGAGGAAGATGAGTATGAAGATCA 4080
Db 8020 TAGTATGGAATTTGGAAGGGGTCAATGTTACTCAATGAGGAAGATGAGTATGAAGATCA 8079
QY 4081 GGCCCTTCTTCCCAAGCAATAGTCCCAATAGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGG 4140
Db 8080 GGCCCTTCTTCCCAAGCAATAGTCCCAATAGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGG 8139
QY 4141 AACATTTTGTGACATTTATGTTATTTGGAATTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 4200
Db 8140 AACATTTTGTGACATTTATGTTATTTGGAATTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 8199

/*tag= b
/standard_name= "single nucleotide polymorphism (SNP)"

QY	4201	CTCAAGCAGACGTGAGTGTGACTTTTGGAAAAAGTATAGGATAAATTAATTAATAATAG	4260
Db	8200	CTCAAGCAGACGTGAGTGTGACTTTTGGAAAAAGTATAGGATAAATTAATTAATAATAG	8259
QY	4261	CACATGATTTCTTTTGTGTTTCTTCAATTTCTGCTACCCCAAGAAAGTAAACAAAGTAT	4320
Db	8260	CACATGATTTCTTTTGTGTTTCTTCAATTTCTGCTACCCCAAGAAAGTAAACAAAGTAT	8319
QY	4321	AGTTTTCAGACAGTTCGTTGTTCAATAATTTCAAGTTCTAGTTGATTCGAGAAATTTTCAAAAT	4380
Db	8320	AGTTTTCAGACAGTTCGTTGTTCAATAATTTCAAGTTCTAGTTGATTCGAGAAATTTTCAAAAT	8379
QY	4381	AAGGAAGAGGGTCTTTTATCTTGTGCTAGGAAAAACCATGACGGAAGGAAAACTGAT	4440
Db	8380	AAGGAAGAGGGTCTTTTATCTTGTGCTAGGAAAAACCATGACGGAAGGAAAACTGAT	8439
QY	4441	GTTTAAAGTCCACTTTTAAACCTATATTTATTTAGTAGGATCTGTCAAGAAAACTTC	4500
Db	8440	GTTTAAAGTCCACTTTTAAACCTATATTTATTTAGTAGGATCTGTCAAGAAAACTTC	8499
QY	4501	CAAAAGATTTAATAATTAACACGACTCTGTTGCAATAAGTTAATGTTTCTGTTTGTG	4560
Db	8500	CAAAAGATTTAATAATTAACACGACTCTGTTGCAATAAGTTAATGTTTCTGTTTGTG	8559
QY	4561	TAATCCACACATTCATCAATGAGTTAGGCTTTGCACTTTGTAAGGAAGGAGGTTCAACAC	4620
Db	8560	TAATCCACACATTCATCAATGAGTTAGGCTTTGCACTTTGTAAGGAAGGAGGTTCAACAC	8619
QY	4621	CTCAATAGCTAATAAACCGTCTTGAATATTTGAAGATTTAAATCTGACTTAGGACG	4680
Db	8620	CTCAATAGCTAATAAACCGTCTTGAATATTTGAAGATTTAAATCTGACTTAGGACG	8679
QY	4681	GGCAGGCTGCTCAGACTATAATCCCAACACTTTGGGAGGCTCAGGCGGCGGTCAACAA	4740
Db	8680	GGCAGGCTGCTCAGACTATAATCCCAACACTTTGGGAGGCTCAGGCGGCGGTCAACAA	8739
QY	4741	GGTCAGGAGTTCAAGACACGCTGACCAATATGTTGGAACCCCATCTCTACTAAATAATAC	4800
Db	8740	GGTCAGGAGTTCAAGACACGCTGACCAATATGTTGGAACCCCATCTCTACTAAATAATAC	8799
QY	4801	AAAAATTAGCAGGCGTGTGTCAGTCTGCTAGGTCCTCCAGCTAGCTGTGAGTGGAG	4860
Db	8800	AAAAATTAGCAGGCGTGTGTCAGTCTGCTAGGTCCTCCAGCTAGCTGTGAGTGGAG	8859
QY	4861	ATTGCATTGAGCCAAAGATC	4879
Db	8860	ATTGCATTGAGCCAAAGATC	8878
RESULT 3			
ID	AAD55117	standard; DNA; 8878 BP.	
XX	AC		
XX	AD55117;		
DT	07-AUG-2003	(first entry)	
XX	Human FBG reference DNA (GI 182597).		
XX	Thrombospondin 2; THBS2; angiotensin converting enzyme; polymorphism;		
KW	ACE-1; beta-fibrinogen; FGB; peripheral vascular disease; ischaemia;		
KW	vascular disease; myocardial infarction; pulmonary embolism; stroke;		
KW	atherosclerosis; coronary artery disease; venous thromboembolism; human;		
KW	Gene; ds.		
XX	Homo sapiens.		
XX	Key	Location/Qualifiers	
FT	variation	replace(5119,T)	
FT	FT	/*tag= a	
FT	FT	/standard_name= "single nucleotide polymorphism (SNP)"	
FT	FT	replace (8059,A)	

FT	WT0003020118-A2.		
XX	13-MAR-2003.		
XX	04-SEP-2002; 2002WO-US028113.		
XX	05-SEP-2001; 2001US-0317178P.		
XX	16-OCT-2001; 2001US-032958P.		
XX	14-DEC-2001; 2001US-00017724.		
XX	(VITI-) VITIVITY INC.		
XX	Mccarthy J;		
XX	WPI; 2003-300816/29.		
XX	Identifying polymorphisms in thrombospondin 2, angiotensin converting enzyme and/or beta-fibrinogen genes in nucleic acid sample of subject, by contacting the nucleic acid with a complementary probe or primer.		
XX	Claim 4; Fig 5; 194pp; English.		
XX	The invention relates to a method for determining the identity of one or more allelic variants of a polymorphic region of a thrombospondin 2 (THBS2), angiotensin converting enzyme (ACE)-1 and/or beta-fibrinogen (FGB) genes in a nucleic acid obtained from a subject. The method involves contacting the nucleic acid with a complementary probe or primer. The method is useful for diagnosing or aiding in the diagnosis of vascular disease or disorder in a subject e.g. myocardial infarction, coronary artery disease, atherosclerosis, ischaemia, stroke, peripheral vascular disease, venous thromboembolism and pulmonary embolism. The present sequence is human FBG reference DNA. Note: This sequence is said to encode SEQ ID NO: 6 (AAB36413). However this does not appear to be the case		
XX	Sequence 8878 BP; 2912 A; 1478 C; 1663 G; 2825 T; 0 U; 0 Other;		
SQ	Query Match 100.0%; Score 4879; DB 9; Length 8878;		
	Best Local Similarity 100.0%; Pred. No. 0;		
	Matches 4879; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
QY	1 TCAGAAATAGTGTTCATGAGTTAAATAACAATGTGGAAGCTGTTCCAGACCTCTCTT	60	
Db	4000 TCAGAAATAGTGTTCATGAGTTAAATAACAATGTGGAAGCTGTTCCAGACCTCTCTT	4059	
QY	61 CTTCTTTTCAGTACATGATTTGCTGAAGACCTCTGGCAAGAGGCGAAGCAAGTAA	120	
Db	4060 CTTCTTTTCAGTACATGATTTGCTGAAGACCTCTGGCAAGAGGCGAAGCAAGTAA	4119	
QY	121 AAGGTAGATATCTTGTGCTTTCCATTCGATTTTCAGCTATAAAATTTGAAACCGTTAGAC	180	
Db	4120 AAGGTAGATATCTTGTGCTTTCCATTCGATTTTCAGCTATAAAATTTGAAACCGTTAGAC	4179	
QY	181 TGCACGAGATGCAATGCTGTGTCAGAAAGATTAACATTTCTGGTGTAGTGAATGATTC	240	
Db	4180 TGCACGAGATGCAATGCTGTGTCAGAAAGATTAACATTTCTGGTGTAGTGAATGATTC	4239	
QY	241 TAGCTTTTGGGACCTTCCCTGGCAATTCGCCAGATAGCACTATTTCAGCTCTTATTC	300	
Db	4240 TAGCTTTTGGGACCTTCCCTGGCAATTCGCCAGATAGCACTATTTCAGCTCTTATTC	4299	
QY	301 CAGTCTGACATCAGCAAGTGTGATTTTCTATGAAAAATTTCTACTAGTCTTATTTTA	360	
Db	4300 CAGTCTGACATCAGCAAGTGTGATTTTCTATGAAAAATTTCTACTAGTCTTATTTTA	4359	
QY	361 AGTATACAAGAACTTGTGACTCAGAAAGATAATTTTACAGGTGGAAGAAAAACCCCTAG	420	
Db	4360 AGTATACAAGAACTTGTGACTCAGAAAGATAATTTTACAGGTGGAAGAAAAACCCCTAG	4419	
QY	421 CATTTATAGTTTTTAACATTTGAGGTTTTGAATGAGAGATTATCCATAATATATTCATTT	480	

[illegible]

Db	5500	GTGTGCATCTGCCTACTGCCTTTGAA	CAGATGGCAGAACCA	CAGATAGATGTCAAGATTTC	5555
QY	1561	TCCAATTTTGTTGTTCTGGAGCTCAT	TGGATATTC	CCAGACACAAAAGTGGAGAAGAC	1620
Db	5560	TCCAATTTTGTTGTTCTGGAGCTCAT	TGGATATTC	CCAGACACAAAAGTGGAGAAGAC	5619
QY	1621	TTTGTTCATCCTCTTAGCAGATAAA	CGCTCC	CAAACTGGGTGGACTTACTAAAGTAA	1680
Db	5620	TTTGTTCATCCTCTTAGCAGATAAA	CGCTCC	CAAACTGGGTGGACTTACTAAAGTAA	5679
QY	1681	ATGAAATCTAAATATTTGTTTATAT	TATATTTT	CAAAGTCTATAATAACACACCTCTTAGTA	1740
Db	5680	ATGAAATCTAAATATTTGTTTATAT	TATATTTT	CAAAGTCTATAATAACACACCTCTTAGTA	5739
QY	1741	ACTTATGTAATGTTTATTTTAAAGAA	TTCGTGAC	TAAATACAAAGTAATATGTCATAAC	1800
Db	5740	ACTTATGTAATGTTTATTTTAAAGAA	TTCGTGAC	TAAATACAAAGTAATATGTCATAAC	5799
QY	1801	CCCTGAAACATAATGTGTCTTACAT	TTTGCAGAA	TGTGAGGAAATATCAGGAAAGGAGT	1860
Db	5800	CCCTGAAACATAATGTGTCTTACAT	TTTGCAGAA	TGTGAGGAAATATCAGGAAAGGAGT	5859
QY	1861	GAAACATCTGAAATGTATCTCAAT	CTCAACCTGAC	AGTCTGTCAAACCGTATAGAGTATAC	1920
Db	5860	GAAACATCTGAAATGTATCTCAAT	CTCAACCTGAC	AGTCTGTCAAACCGTATAGAGTATAC	5919
QY	1921	TGTGACATGAATACAGAAAATGGAG	TAAGCTTT	CGACAGTTGTGACCTGTTGATCTGT	1980
Db	5920	TGTGACATGAATACAGAAAATGGAG	TAAGCTTT	CGACAGTTGTGACCTGTTGATCTGT	5979
QY	1981	AATTAATTTGGATACCGTAAATG	CCAGAAAC	CAAGCCAGGTGTGGTGCTCATACCTGT	2040
Db	5980	AATTAATTTGGATACCGTAAATG	CCAGAAAC	CAAGCCAGGTGTGGTGCTCATACCTGT	6039
QY	2041	AAITTCAGACCTTTGGAGCCAAAG	TGGGCTGAT	CTTGAGCCTAGGAGTTTGAACCT	2100
Db	6040	AAITTCAGACCTTTGGAGCCAAAG	TGGGCTGAT	CTTGAGCCTAGGAGTTTGAACCT	6099
QY	2101	AGCCTGGGCAACATAATGAGACCT	TAACCTCA	CAAAAATAAATAATACCAAAAAAAA	2160
Db	6100	AGCCTGGGCAACATAATGAGACCT	TAACCTCA	CAAAAATAAATAATACCAAAAAAAA	6159
QY	2161	AAAAAATCAGCTGTGTGGTAGTAT	GTGCTGTAGT	TCCAGCTATCCAGAGGCTGAGA	2220
Db	6160	AAAAAATCAGCTGTGTGGTAGTAT	GTGCTGTAGT	TCCAGCTATCCAGAGGCTGAGA	6219
QY	2221	TGGGAGTACCTGAGCCCAACCTGG	AGTCTTGAT	CACTGAACTGTAGCCTGG	2280
Db	6220	TGGGAGTACCTGAGCCCAACCTGG	AGTCTTGAT	CACTGAACTGTAGCCTGG	6279
QY	2281	GCAACAGAGGATAGTGAGATCTGT	CTCAAAA	AAAAAATAATTAATTAAGCCAGGAAC	2340
Db	6280	GCAACAGAGGATAGTGAGATCTGT	CTCAAAA	AAAAAATAATTAATTAAGCCAGGAAC	6339
QY	2341	AAGACTTAGCTTAACATCTAA	CATAGCTG	CAAAAGGAGTAATTTGATGTGAATTCAC	2400
Db	6340	AAGACTTAGCTTAACATCTAA	CATAGCTG	CAAAAGGAGTAATTTGATGTGAATTCAC	6399
QY	2401	CTGATATTTAAAGTTATAAATAT	CTATATTC	CAATTTGGGTAAGATAAAGCACTT	2460
Db	6400	CTGATATTTAAAGTTATAAATAT	CTATATTC	CAATTTGGGTAAGATAAAGCACTT	6459
QY	2461	GCAGTTTCCAAAGATTTTACAAG	TTTTTAC	CTCTCATATTTATTTCCCTATTGTCATTTT	2520
Db	6460	GCAGTTTCCAAAGATTTTACAAG	TTTTTAC	CTCTCATATTTATTTCCCTATTGTCATTTT	6519
QY	2521	TAGAGCCAAATATATACTAAAT	TGGAATG	GACAGGGGATTCAGATATATTTTCAAAGT	2580
Db	6520	TAGAGCCAAATATATACTAAAT	TGGAATG	GACAGGGGATTCAGATATATTTTCAAAGT	6579
QY	2581	GACATTAATTTGCTGTTGGTTAA	TATATATG	CTCTTTTGTGTCACCAAGGATGAC	2640
Db	6580	GACATTAATTTGCTGTTGGTTAA	TATATATG	CTCTTTTGTGTCACCAAGGATGAC	6639

QY 2641 AGTGATTCAGAACCGTCAAGACGGTAGTGTGACCTTTGGCAGGAATGGGATCCATATAA 2700
DB |||||||
QY 6640 AGTGATTCAGAACCGTCAAGACGGTAGTGTGACCTTTGGCAGGAATGGGATCCATATAA 6699
DB |||||||
QY 2701 ACAGGGATTTGGAATGTTTGAACCAACACAGATGGGAAGAAATTAAGTGTGGCTTACCCAGG 2760
DB |||||||
QY 6700 ACAGGGATTTGGAATGTTTGAACCAACACAGATGGGAAGAAATTAAGTGTGGCTTACCCAGG 6759
DB |||||||
QY 2761 TAACGAAACAGGCATGCAAAATAAAATCAATCTTATTGAAATGGGATTTTTTTTAAATAAA 2820
DB |||||||
QY 6760 TAACGAAACAGGCATGCAAAATAAAATCAATCTTATTGAAATGGGATTTTTTTTAAATAAA 6819
DB |||||||
QY 2821 AAAACATTCAATGTTGGAACCTGTTTAGGCACTTAAAGAGGAGTTTCCCTGACAAAATGCT 2880
DB |||||||
QY 6820 AAAACATTCAATGTTGGAACCTGTTTAGGCACTTAAAGAGGAGTTTCCCTGACAAAATGCT 6879
DB |||||||
QY 2881 GGAAGCTAAAGATAAGGGAAGAAAGGAGTCTTTTGGTCTTCCAAAATTTTTTTTGGTG 2940
DB |||||||
QY 6880 GGAAGCTAAAGATAAGGGAAGAAAGGAGTCTTTTGGTCTTCCAAAATTTTTTTTGGTG 6939
DB |||||||
QY 2941 AGAGATTTTTATTGTTTCTTTTGGTGAATTAATTTGGCTTGGAAATGATAAAATTAGCC 3000
DB |||||||
QY 6940 AGAGATTTTTATTGTTTCTTTTGGTGAATTAATTTGGCTTGGAAATGATAAAATTAGCC 6999
DB |||||||
QY 3001 AGCTTACAGAGTGGACCCACAGAACTTTTGNATGAATGGAGGACTGGAAGGAGACA 3060
DB |||||||
QY 7000 AGCTTACAGAGTGGACCCACAGAACTTTTGNATGAATGGAGGACTGGAAGGAGACA 7059
DB |||||||
QY 3061 AAGTAAAGCTCACTATGAGGAGTCACTGTACAGAAATGAAGCAACAAATACAGATCT 3120
DB |||||||
QY 7060 AAGTAAAGCTCACTATGAGGAGTCACTGTACAGAAATGAAGCAACAAATACAGATCT 7119
DB |||||||
QY 3121 CAGTGAACAAATACAGAGGAACAGCCGGTAAATGCCCTCATGAGTGGAGCATCTCAGCTGA 3180
DB |||||||
QY 7120 CAGTGAACAAATACAGAGGAACAGCCGGTAAATGCCCTCATGAGTGGAGCATCTCAGCTGA 7179
DB |||||||
QY 3181 TGGGAGAAACAGGACCATGACCACTTCAACAGGCATGTTCTTCAGCAGCTATGACAGAG 3240
DB |||||||
QY 7180 TGGGAGAAACAGGACCATGACCACTTCAACAGGCATGTTCTTCAGCAGCTATGACAGAG 7239
DB |||||||
QY 3241 ACAATGAGGCTGTATGTGTGACCTTCTTGTCTGCTTTTAAATACACACTAATATC 3300
DB |||||||
QY 7240 ACAATGAGGCTGTATGTGTGACCTTCTTGTCTGCTTTTAAATACACACTAATATC 7299
DB |||||||
QY 3301 ATTACTCAGAAATCAATTAACAATAATTTTAAATAGCTACCACTTCTGGGCACTTACTGTCA 3360
DB |||||||
QY 7300 ATTACTCAGAAATCAATTAACAATAATTTTAAATAGCTACCACTTCTGGGCACTTACTGTCA 7359
DB |||||||
QY 3361 GCCACTGTCTAAGCTCTTTATGATCAGCTTCAAGAGCACTTCACTATAGGTTAGACATT 3420
DB |||||||
QY 7360 GCCACTGTCTAAGCTCTTTATGATCAGCTTCAAGAGCACTTCACTATAGGTTAGACATT 7419
DB |||||||
QY 3421 CTTATTCTCAATTTTACAGATGAGATTTAGAGAGATTACGTGATTGTCCTCAATGTACACA 3480
DB |||||||
QY 7420 CTTATTCTCAATTTTACAGATGAGATTTAGAGAGATTACGTGATTGTCCTCAATGTACACA 7479
DB |||||||
QY 3481 ACTACCAGAGATAAAACTAGAAATTTGAGCAAGTCTTCTGAAATAATGAGCATTTAG 3540
DB |||||||
QY 7480 ACTACCAGAGATAAAACTAGAAATTTGAGCAAGTCTTCTGAAATAATGAGCATTTAG 7539
DB |||||||
QY 3541 ATAAATACCTATATCTCTATATTTCTAAAGTGTGTGAAACTTTTCATTTCCAG 3600
DB |||||||
QY 7540 ATAAATACCTATATCTCTATATTTCTAAAGTGTGTGAAACTTTTCATTTCCAG 7599
DB |||||||
QY 3601 GGTTCCTGATACCTAAGGGTTGTAAGCTAATTTCCAGTATAAAGTAAACAAACAGT 3660
DB |||||||
QY 7600 GGTTCCTGATACCTAAGGGTTGTAAGCTAATTTCCAGTATAAAGTAAACAAACAGT 7659
DB |||||||
QY 3661 CCCTAGATGGATTGCCAAGAGGCCAGTTATCTCTCTTTCTGCTATAGGGCAGGAG 3720
DB |||||||
QY 7660 CCCTAGATGGATTGCCAAGAGGCCAGTTATCTCTCTTTCTTGTATAGGGCAGGAG 7719
DB |||||||

QY 3721 GTCTTTGGTGTATTAGTGTGACTCTATGTATAGCACCCAAAGAAAGAACTACTGTGCACA 3780
DB |||||||
QY 7720 GTCTTTGGTGTATTAGTGTGACTCTATGTATAGCACCCAAAGAAAGAACTACTGTGCACA 7779
DB |||||||
QY 3781 CGAGGTGAGCAGTCTTTTATGGGTAATCTGMAAACGTAACTTGACCAACCGTGTCTGT 3840
DB |||||||
QY 7780 CGAGGTGAGCAGTCTTTTATGGGTAATCTGCAAAAACGTAACTTGACCAACCGTGTCTGT 7839
DB |||||||
QY 3841 TTCTAATAACGCCAAACACATTTCTTTTCAGGTAAACATCAGATCCAGAAAACAGTGT 3900
DB |||||||
QY 7840 TTCTAATAACGCCAAACACATTTCTTTTCAGGTAAACATCAGATCCAGAAAACAGTGT 7899
DB |||||||
QY 3901 CTAAGAAGACGGTGGTGGTATPAAATAGATGTCATCCAGCAATCCAAACGGCA 3960
DB |||||||
QY 7900 CTAAGAAGACGGTGGTGGTATPAAATAGATGTCATCCAGCAATCCAAACGGCA 7959
DB |||||||
QY 3961 GATATCTCTGGGGTGGACAGTACATCTGGGACATCGCAAGCATGGCAGAGATGATGGTG 4020
DB |||||||
QY 7960 GATATCTCTGGGGTGGACAGTACATCTGGGACATCGCAAGCATGGCAGAGATGATGGTG 8019
DB |||||||
QY 4021 TAGTATGGATGAATTTGGAAGGGGTCTATGTACTCTCAATGAGGAAGATGAGTATGAAGATCA 4080
DB |||||||
QY 8020 TAGTATGGATGAATTTGGAAGGGGTCTATGTACTCTCAATGAGGAAGATGAGTATGAAGATCA 8079
DB |||||||
QY 4081 GGCCCTCTTCCACAGCAATAGTCCCAATACGTAGATTTTGTCTCTCTGTATGTGAC 4140
DB |||||||
QY 8080 GGCCCTCTTCCACAGCAATAGTCCCAATACGTAGATTTTGTCTCTCTGTATGTGAC 8139
DB |||||||
QY 4141 AACATTTTGTACATTTATTTGGAATTTCTTTTCATACATTTATTTTCTCTTAAACT 4200
DB |||||||
QY 8140 AACATTTTGTACATTTATTTGGAATTTCTTTTCATACATTTATTTTCTCTTAAACT 8199
DB |||||||
QY 4201 CTCACGACAGCTGAGTGTGACTTTTGAAGAAAGTATAGGATAAAATTCATTTAAATAG 4260
DB |||||||
QY 8200 CTCACGACAGCTGAGTGTGACTTTTGAAGAAAGTATAGGATAAAATTCATTTAAATAG 8259
DB |||||||
QY 4261 CACATGATTTCTTTTGTCTTTCTTCATTTCTCTGTCTACCCAAAGAAAGTAAACAAAGTAT 4320
DB |||||||
QY 8260 CACATGATTTCTTTTGTCTTTCTTCATTTCTCTGTCTACCCAAAGAAAGTAAACAAAGTAT 8319
DB |||||||
QY 4321 AGTTTTGACAGAGTGGTGTTCATATTTTCAGTTCTAGTTGATTCGAGNATTTTCAAT 4380
DB |||||||
QY 8320 AGTTTTGACAGAGTGGTGTTCATATTTTCAGTTCTAGTTGATTCGAGNATTTTCAAT 8379
DB |||||||
QY 4381 AAGGAAGAGGGTCTTTTATCTTGTGTAAGAAACCATGACGGAAGAAAGAAACTGAT 4440
DB |||||||
QY 8380 AAGGAAGAGGGTCTTTTATCTTGTGTAAGAAACCATGACGGAAGAAAGAAACTGAT 8439
DB |||||||
QY 4441 GTTTAAAGAGTCCACTTTTAAACTATATTTTATTTATGTAGGATCTGTCAAAGAAACTTC 4500
DB |||||||
QY 8440 GTTTAAAGAGTCCACTTTTAAACTATATTTTATTTATGTAGGATCTGTCAAAGAAACTTC 8499
DB |||||||
QY 4501 CAAAAAGATTTTATTAATTAACCCAGACTCTGTTGCAATAGTTAAATGTTTCTTGTGTTTG 4560
DB |||||||
QY 8500 CAAAAAGATTTTATTAATTAACCCAGACTCTGTTGCAATAGTTAAATGTTTCTTGTGTTTG 8559
DB |||||||
QY 4561 TAATCCACACATTTCAATGAGTTAGGCTTTGCACCTTGTAAAGAAAGGAAAGCTTCACAAC 4620
DB |||||||
QY 8560 TAATCCACACATTTCAATGAGTTAGGCTTTGCACCTTGTAAAGAAAGGAAAGCTTCACAAC 8619
DB |||||||
QY 4621 CTCAAATAGCTAATAAACCCGGTCTTGAATATTTGAAGATTTAAATCTGATCTAGGAGC 4680
DB |||||||
QY 8620 CTCAAATAGCTAATAAACCCGGTCTTGAATATTTGAAGATTTAAATCTGATCTAGGAGC 8679
DB |||||||
QY 4681 GGCAAGGTGTCTCAAGCTATAATCCCAACAATTTTGGAGGCTGAGCGGGCGGTCTCAA 4740
DB |||||||
QY 8680 GGCAAGGTGTCTCAAGCTATAATCCCAACAATTTTGGAGGCTGAGCGGGCGGTCTCAA 8739
DB |||||||
QY 4741 GGTGAGAGTTCAAGACAGCCTGACCAATATGTTGAAGAAACCCCATCTCTACTAAAAATAC 4800
DB |||||||
QY 8740 GGTGAGAGTTCAAGACAGCCTGACCAATATGTTGAAGAAACCCCATCTCTACTAAAAATAC 8799
DB |||||||
QY 4801 AAAAAATTAGCAGCGGTGGTGGCAGGTGCCTGTAGGTCCAGCTAGCCTGTGAGGTGGAG 4860
DB |||||||

DB 8800 AAAAAATAGCCAGCGTGTGGCAGGTGCTGTAGGTCCAGCTAGCTGTGAGTGGAG 8859
QY 4861 ATTGCATTGAGCCCAAGATC 4879
DB 8860 ATTGCATTGAGCCCAAGATC 8878
RESULT 4
ADP65553
ID ADP65553 standard; DNA; 8878 BP.
XX AC ADP65553;
XX DT 12-AUG-2004 (first entry)
XX DE Human fibrinogen beta chain (FGB), complete cds DNA.
XX KW autoimmune disease; arthritis; gene expression analysis;
KW rheumatoid arthritis; collagen-induced; immunosuppressive; antirheumatic;
KW antiarthritic; osteopathic; antigout; antinflammatory; dermatological;
KW immunomodulatory; lupus; ankylosing spondylitis; Fibrositis;
KW fibromyalgia; osteoarthritis; gout; juvenile rheumatoid arthritis;
KW immune; ds; human.
XX OS Homo sapiens.
XX XN WO2003072827-A1.
XX PN 04-SEP-2003.
XX PD 31-OCT-2002; 2002WO-US035433.
XX PF 31-OCT-2001; 2001US-0336220P.
XX PR (CHIL-) CHILDREN'S HOSPITAL MEDICAL CENT.
XX PA Hirsch R, Thorton SL;
XX PI WPI; 2003-712740/67.
XX DR GENBANK; M64983.
XX PT Diagnosing and analyzing autoimmune disease using gene expression
PT profiles and microarray technology, useful for diagnosing and treating
PT rheumatoid arthritis, lupus, fibrositis, osteoarthritis, fibromyalgia and
PT gout.
XX PS Disclosure; Page; 56pp; English.
XX CC The invention relates to a novel method for diagnosing and analysing
CC autoimmune disease or arthritides. The method comprises obtaining a
CC patient sample containing mRNA, analysing gene expression using the mRNA
CC that results in a gene expression signature of the mRNA, and using that
CC gene expression signature to diagnose or analyse the autoimmune disease
CC or arthritides in the patient, where gene expression of at least 60% of
CC the genes correlates with that of the gene signature. The invention
CC further comprises: a treatment of rheumatoid arthritis; identification of
CC genes for targeting in the treatment of rheumatoid arthritis in a mammal
CC other than a mouse; diagnosis of rheumatoid arthritis in a mammal; an
CC array or gene chip, specific for rheumatoid arthritis; diagnosis or
CC analyses of autoimmune disease or rheumatoid arthritis; screening the
CC efficacy of a candidate drug in vitro for the treatment of collagen-
CC induced arthritis; and reducing the symptoms associated with collagen-
CC induced arthritis. The compositions of the invention have the following
CC activities: immunosuppressive, antirheumatic, antiarthritic, osteopathic,
CC antigout, antinflammatory, dermatological, and immunomodulatory. The
CC methods and compositions of the present invention are useful for
CC diagnosing and treating autoimmune disease or arthritides, such as
CC rheumatoid arthritis, lupus, ankylosing spondylitis, fibrositis,
CC fibromyalgia, osteoarthritis, gout, juvenile rheumatoid arthritis, and an
CC immune disease caused by an infectious agent. This polynucleotide
CC represents a DNA sequence relating to the genes used in the analysis and
CC treatment of autoimmune diseases or arthritides. Note: This sequence is

CC not shown in the specification. It has been supplied in an electronic
CC format from WIFO.

SQ Sequence 8878 BP; 2912 A; 1478 C; 1663 G; 2825 T; 0 U; 0 Other;
XX

Query Match 100.0%; Score 4879; DB 11; Length 8878;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 4879; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCAGAAATAGTGTGATGAGTTAAATAACAATGTGGAAGCTGTTCCCGAGACCTCTCTT 60
DB 4000 TCAGAAATAGTGTGATGAGTTAAATAACAATGTGGAAGCTGTTCCCGAGACCTCTCTT 4059

QY 61 CTTCCCTTCAGTACATGTATTTGCTGAAAGACCTGTGCAAAAGAGGCGAGCAAGTAA 120
DB 4060 CTTCCCTTCAGTACATGTATTTGCTGAAAGACCTGTGCAAAAGAGGCGAGCAAGTAA 4119

QY 121 AAGGTAGATATCTGTGTGCTTTCCATTCCATTTCAGCTTATAAAATTTGGAACCGTTAGAC 180
DB 4120 AAGGTAGATATCTGTGTGCTTTCCATTTCAGCTTATAAAATTTGGAACCGTTAGAC 4179

QY 181 TGCCACGAGAATGCATGTTGTGAGAGATTAAACATTTCTGGGTAGTGAATGACATTC 240
DB 4180 TGCCACGAGAATGCATGTTGTGAGAGATTAAACATTTCTGGGTAGTGAATGACATTC 4239

QY 241 TACGGCTTTTGGGCACCTTCCCTGCAACTTTCAGATAGCAGTATTCAGCTCTTATTC 300
DB 4240 TACGGCTTTTGGGCACCTTCCCTGCAACTTTCAGATAGCAGTATTCAGCTCTTATTC 4299

QY 301 CAGTCTGACATCAGCAAGTGTGATTTCTATGAAAAATTTCTACTATGATCTCTTATTTA 360
DB 4300 CAGTCTGACATCAGCAAGTGTGATTTCTATGAAAAATTTCTACTATGATCTCTTATTTA 4359

QY 361 AGTATACAGAACTTTGTGACTCAGAGATAATATTACAGAGTGGAAAAAACCCTTAG 420
DB 4360 AGTATACAGAACTTTGTGACTCAGAGATAATATTACAGAGTGGAAAAAACCCTTAG 4419

QY 421 CATTTATAGTTTAAACATTTGAGGTTTGAATGAGAGATTATTCATATATATATTTCA 480
DB 4420 CATTTATAGTTTAAACATTTGAGGTTTGAATGAGAGATTATTCATATATATATTTCA 4479

QY 481 GTGTGTGGATAATGACACCTAACTGTGAATTTTGTAGGTGAGAATTTGTAGTGTGTTG 540
DB 4480 GTGTGTGGATAATGACACCTAACTGTGAATTTTGTAGGTGAGAATTTGTAGTGTGTTG 4539

QY 541 ACTTGTGCTCAGGAAACAGCTAGTGGCTGGCCGACAGGCACTCTCAGTGTAGTACA 600
DB 4540 ACTTGTGCTCAGGAAACAGCTAGTGGCTGGCCGACAGGCACTCTCAGTGTAGTACA 4599

QY 601 TACCCACAGTTGGAAAAATTTTCAAAGAAATCAAAGGAATCATGACATCTTATAAATTTCA 660
DB 4600 TACCCACAGTTGGAAAAATTTTCAAAGAAATCAAAGGAATCATGACATCTTATAAATTTCA 4659

QY 661 AGGTTCTGCTATCTATGTGAAATTTGGAATAAATCAAGCATATTCACCTCTGTAGAT 720
DB 4660 AGGTTCTGCTATCTATGTGAAATTTGGAATAAATCAAGCATATTCACCTCTGTAGAT 4719

QY 721 TGAACCTCTCAGATGGAGACCCCAATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 780
DB 4720 TGAACCTCTCAGATGGAGACCCCAATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 4779

QY 781 TAAACAACCTATTTTCAATTTTACTGGACACAACTTTTAGCGTATACCTATGGTAAATTA 840
DB 4780 TAAACAACCTATTTTCAATTTTACTGGACACAACTTTTAGCGTATACCTATGGTAAATTA 4839

QY 841 CTAGTATGGTGGTTAGGATTTATGTTAATTTGTATATGTCAAGCCCAAAATCATTTCCAC 900
DB 4840 CTAGTATGGTGGTTAGGATTTATGTTAATTTGTATATGTCAAGCCCAAAATCATTTCCAC 4899

QY 901 TAAATATGACTAT 960
DB 4900 TAAATATGACTAT 4959

QY 961 AGAATCAAAATTTGATAGTTAAATACATTAAGTTTATGAGGCCAAAATGCTAACTATT 1020
DB |||||
QY 4960 AGAAAATCAAAATTTGATAGTTAAATACATTAAGTTTATGAGGCCAAAATGCTAACTATT 5019
DB |||||
QY 1021 TCTACATAATTTCAATTTTCCAGATAATGAAATGAGTCAATGAGTACTCTCGAAGT 1080
DB |||||
QY 5020 TCTACATAATTTCAATTTTCCAGATAATGAAATGAGTCAATGAGTACTCTCGAAGT 5079
DB |||||
QY 1081 GGAAAGCAACCAATTAATATATAGATGAGACTGTGAATAGCAATATCCCAACTAACTTCG 1140
DB |||||
QY 5080 GGAAAGCAACCAATTAATATATAGATGAGACTGTGAATAGCAATATCCCAACTAACTTCG 5139
DB |||||
QY 1141 TGTGCTTCCTCAATCCTCGMAAACCTGAGAGCAAAATACAAAAGTTAGAAATCTGATGT 1200
DB |||||
QY 5140 TGTGCTTCCTCAATCCTCGMAAACCTGAGAGCAAAATACAAAAGTTAGAAATCTGATGT 5199
DB |||||
QY 1201 CTCAGCTCAAAATGGAATATTTGCGCACCCCATGCACTGTCAAGTTCGAATATTCCTGTGT 1260
DB |||||
QY 5200 CTCAGCTCAAAATGGAATATTTGCGCACCCCATGCACTGTCAAGTTCGAATATTCCTGTGT 5259
DB |||||
QY 1261 GTCTGCAAGGTAATCTGAATCTATAAACAATATTTTATAGAGAGTTCCAGAGAACTCACAC 1320
DB |||||
QY 5260 GTCTGCAAGGTAATCTGAATCTATAAACAATATTTTATAGAGAGTTCCAGAGAACTCACAC 5319
DB |||||
QY 1321 ACCAAAATAAGAGAACAAACAACAACAATAAATCTAAGTGGATTTTCCCAACAGATC 1380
DB |||||
QY 5320 ACCAAAATAAGAGAACAAACAACAACAATAAATCTAAGTGGATTTTCCCAACAGATC 5379
DB |||||
QY 1381 ATAATGACATTAACAGTACATCATAAAAATATCCTTAGCCAGTTGTGTGGAAGTGGCCCT 1440
DB |||||
QY 5380 ATAATGACATTAACAGTACATCATAAAAATATCCTTAGCCAGTTGTGTGGAAGTGGCCCT 5439
DB |||||
QY 1441 GGTGCATTTGCTGGTTTGTATGACGAGGATGGGGCACAGGTAGTCCAGGGTGGCTGAT 1500
DB |||||
QY 5440 GGTGCATTTGCTGGTTTGTATGACGAGGATGGGGCACAGGTAGTCCAGGGTGGCTGAT 5499
DB |||||
QY 1501 GTGTGCATCTGCTACTGCTTCAACAGATGGCAGAACACAGATAGATGACAGGTTTC 1560
DB |||||
QY 5500 GTGTGCATCTGCTACTGCTTCAACAGATGGCAGAACACAGATAGATGACAGGTTTC 5559
DB |||||
QY 1561 TCCATTTTGTGTCTTGGAGTCTATGGATATTTCCAGGACACAAAAGGTGGAGAGAGC 1620
DB |||||
QY 5560 TCCATTTTGTGTCTTGGAGTCTATGGATATTTCCAGGACACAAAAGGTGGAGAGAGC 5619
DB |||||
QY 1621 TTTGTTCACTCTTAGCAGATAAACGTCTCTCAAACTGGGGTGGACTTACTAAAGTAAA 1680
DB |||||
QY 5620 TTTGTTCACTCTTAGCAGATAAACGTCTCTCAAACTGGGGTGGACTTACTAAAGTAAA 5679
DB |||||
QY 1681 ATGAAATCTAATATTTGTTATATTTTCAAGGTCATATAACACACTCCTTAGTA 1740
DB |||||
QY 5680 ATGAAATCTAATATTTGTTATATTTTCAAGGTCATATAACACACTCCTTAGTA 5739
DB |||||
QY 1741 ACTTATGTAATTTATTTTAAAGAAATGGTGACTAAATACAAAAGTAAATATGTCATAAAC 1800
DB |||||
QY 5740 ACTTATGTAATTTATTTTAAAGAAATGGTGACTAAATACAAAAGTAAATATGTCATAAAC 5799
DB |||||
QY 1801 CCTGGAACATAATGTTGCTTCAATTTGCGAGATGTGAGAAATTTATCAGAAAGGAGGT 1860
DB |||||
QY 5800 CCTGGAACATAATGTTGCTTCAATTTGCGAGATGTGAGAAATTTATCAGAAAGGAGGT 5859
DB |||||
QY 1861 GAAACATCTGAAATGTAATCTCAATTTCAACCTGACAGTTCTGTCAAAACCGTATAGAGTATAC 1920
DB |||||
QY 5860 GAAACATCTGAAATGTAATCTCAATTTCAACCTGACAGTTCTGTCAAAACCGTATAGAGTATAC 5919
DB |||||
QY 1921 TGTGACATGAATACAGAAAATGGAGGTAAAGCTTTTCGACAGTTGTTGACCTGTTGATCTGT 1980
DB |||||
QY 5920 TGTGACATGAATACAGAAAATGGAGGTAAAGCTTTTCGACAGTTGTTGACCTGTTGATCTGT 5979
DB |||||
QY 1981 AATTAATTTGGATACCGTAAATGTCAGGAAACAAAGGCCAGGTGGTGGCTCATACCTGT 2040
DB |||||
QY 5980 AATTAATTTGGATACCGTAAATGTCAGGAAACAAAGGCCAGGTGGTGGCTCATACCTGT 6039
DB |||||
QY 2041 AATTCAGCACCTTTGGGAGGCCAAAGTGGGCTGATAGCTTAGGCTTAGGAGTTTGAAC 2100
DB |||||

DB 6040 AATTCAGCACCTTTGGGAGGCCAAAGTGGGCTGATAGCTTAGGCTTAGGAGTTTGAAC 6099
QY |||||
DB 2101 AGCTGGGGCAACATTAATGAGACCTTAACCTCTCAAAAACAAAATACCAAAAAAAA 2160
QY |||||
DB 6100 AGCTGGGGCAACATTAATGAGACCTTAACCTCTCAAAAACAAAATACCAAAAAAAA 6159
QY |||||
QY 2161 AAAAAAATCAGCTGTGTTGGTAGTATGTCCTGTAGTCCAGCTATCCAGGAGGCTGAGA 2220
DB |||||
QY 6160 AAAAAAATCAGCTGTGTTGGTAGTATGTCCTGTAGTCCAGCTATCCAGGAGGCTGAGA 6219
DB |||||
QY 2221 TGGGAGATCACTGAGCCCAACCTGGAGTCTTGATCATGTACTGAACTGTAGGCTGG 2280
DB |||||
QY 6220 TGGGAGATCACTGAGCCCAACCTGGAGTCTTGATCATGTACTGAACTGTAGGCTGG 6279
DB |||||
QY 2281 GCAACAGAGGATAGTGAGATCTGTCTCAAAAACAAAATTAATTAAGCCAGGAAAC 2340
DB |||||
QY 6280 GCAACAGAGGATAGTGAGATCTGTCTCAAAAACAAAATTAATTAAGCCAGGAAAC 6339
DB |||||
QY 2341 AAGACTTAGCTCTAATCTAATCACTAGCTGACAAAGGAGTAAATTTGATGTGGAATCAAC 2400
DB |||||
QY 6340 AAGACTTAGCTCTAATCTAATCACTAGCTGACAAAGGAGTAAATTTGATGTGGAATCAAC 6399
DB |||||
QY 2401 CTGATATTTAAAGTTTATAAATAATCTAATTAATTTCAAAATTTGGGGTAAAGTAAAGCACTT 2460
DB |||||
QY 6400 CTGATATTTAAAGTTTATAAATAATCTAATTAATTTCAAAATTTGGGGTAAAGTAAAGCACTT 6459
DB |||||
QY 2461 GCAGTTTCAAGATTTTACAAGTTTACCTCTCATATTTATTTCTCTTATTTGCTCTATTT 2520
DB |||||
QY 6460 GCAGTTTCAAGATTTTACAAGTTTACCTCTCATATTTATTTTCTTATTTGCTCTATTT 6519
DB |||||
QY 2521 TAGAGCACCAATATATATCTAATGGAATGGACAGGGGATTCAGATATTTATTTCAAGT 2580
DB |||||
QY 6520 TAGAGCACCAATATATATCTAATGGAATGGACAGGGGATTCAGATATTTATTTCAAGT 6579
DB |||||
QY 2581 GACATTAATTTGCTGTTGTTTAAATATATATGCTCTTTTTGTTTCTGTCACAAAGGATGGAC 2640
DB |||||
QY 6580 GACATTAATTTGCTGTTGTTTAAATATATATGCTCTTTTTGTTTCTGTCACAAAGGATGGAC 6639
DB |||||
QY 2641 AGTGATTCAGACCGTCAAGACGAGTGTGATGCTTTGGCAGGAAATGGATCCATATAA 2700
DB |||||
QY 6640 AGTGATTCAGACCGTCAAGACGAGTGTGATGCTTTGGCAGGAAATGGATCCATATAA 6699
DB |||||
QY 2701 ACAGGATTTGGAATTTGCAACCAACACAGATGGGAATTTACTGTGGCTACCAGG 2760
DB |||||
QY 6700 ACAGGATTTGGAATTTGCAACCAACACAGATGGGAATTTACTGTGGCTACCAGG 6759
DB |||||
QY 2761 TAAACGACAGGCAATGCAAAATAAATAATCATTTCTATTTGAAATGGGATTTTTTTTAAATPAAA 2820
DB |||||
QY 6760 TAAACGACAGGCAATGCAAAATAAATAATCATTTCTATTTGAAATGGGATTTTTTTTAAATPAAA 6819
DB |||||
QY 2821 AAACATTTCTATTTGGAAGCTGTTTGGCAGTTTAAAGGAGTTTCTGACAAAATGT 2880
DB |||||
QY 6820 AAACATTTCTATTTGGAAGCTGTTTGGCAGTTTAAAGGAGTTTCTGACAAAATGT 6879
DB |||||
QY 2881 GGAAGCTAAAGATAAGGAAGGAAGGAGGTTTGGTTCCTCCAAAATTTTATTTTGGTG 2940
DB |||||
QY 6880 GGAAGCTAAAGATAAGGAAGGAAGGAGGTTTGGTTCCTCCAAAATTTTATTTTGGTG 6939
DB |||||
QY 2941 AGAGATTTTATTTTGTGTTTCTTTTGGTGAATTAATTTGGCTTGGAAATGATAAATTTAGCC 3000
DB |||||
QY 6940 AGAGATTTTATTTTGTGTTTCTTTTGGTGAATTAATTTGGCTTGGAAATGATAAATTTAGCC 6999
DB |||||
QY 3001 AGCTTACAGGATGGGACCCACAGAACTTTGATAGAAATGGAGGACTGGAAGGAGACA 3060
DB |||||
QY 7000 AGCTTACAGGATGGGACCCACAGAACTTTGATAGAAATGGAGGACTGGAAGGAGACA 7059
DB |||||
QY 3061 AAGTAAAGGCTCACTATGGAGGATTCCTCTCAGAAATGAAGCAACAAATACCAAGTCT 3120
DB |||||
QY 7060 AAGTAAAGGCTCACTATGGAGGATTCCTCTCAGAAATGAAGCAACAAATACCAAGTCT 7119
DB |||||
QY 3121 CAGTGAACAAATACAGAGGAACAGCCGTTAATGCCCTCATGGATGGAGCATCTCAGCTGA 3180
DB |||||

(CHIL-) CHILDREN'S HOSPITAL MEDICAL CENT.

Hirsch R, Thorton SL;

WPI; 2003-712740/67.

GENBANK; M64983.

Diagnosing and analyzing autoimmune disease using gene expression

profiles and microarray technology, useful for diagnosing and treating rheumatoid arthritis, lupus, fibrositis, osteoarthritis, fibromyalgia and gout.

•
•
•
•
•

Disclosure; Page; 56pp; English.

The invention relates to a novel method for diagnosing and analysing

autoimmune disease or arthritis). The method comprises obtaining a patient sample containing mRNA, analysing gene expression using the mRNA that results in a gene expression signature of the mRNA, and using that gene expression signature to diagnose or analyse the autoimmune disease or arthritis in the patient, where gene expression of at least 60% of the genes correlates with that of the gene signature. The invention further comprises: a treatment of rheumatoid arthritis; identification of genes for targeting in the treatment of rheumatoid arthritis in a mammal other than a mouse; diagnosis of rheumatoid arthritis in a mammal; an array or gene chip, specific for rheumatoid arthritis; diagnosis or analyses of autoimmune disease or rheumatoid arthritis; screening the efficacy of a candidate drug *in vitro* for the treatment of collagen-induced arthritis; and reducing the symptoms associated with collagen-induced arthritis. The compositions of the invention have the following activities: immunosuppressive, antirheumatic, antiarthritic, osteopathic, antigout, antiinflammatory, dermatological, and immunomodulatory. The methods and compositions of the present invention are useful for diagnosing and treating autoimmune disease or arthritis, such as rheumatoid arthritis, lupus, ankylosing spondylitis, fibrositis, fibromyalgia, osteoarthritis, gout, juvenile rheumatoid arthritis, and an immune disease caused by an infectious agent. This polynucleotide represents a DNA sequence relating to the genes used in the analysis and treatment of autoimmune diseases or arthritis. Note: this sequence is not shown in the specification. It has been supplied in an electronic format from WIPO.

Sequence 8878 BP: 2912 A: 1478 C: 1663 G: 2825 T: 0 U: 0 Other:

every Match 100.0%; Score 4879; DB 11; Length 8878;

```

Very Match      score 4875; DB 11; length 6095;
Best Local Similarity    100.0%; Pred. No. 0;

```

Matches 4879; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 TCAGAAATAGTGTGATGAGTTAAATAACAAATGTGGAAGCTGTTTCCAGACCTCCTCTT 60

4000 TCAGAAATAGTGTGATGAGTTAAATAACAAATGTGGAAGCTGTTTCCAGACCTCCTCTT 4059

61 CTTCTTTCAGTACATGTATTTGCTGAAAGACCTGTGGCAAAGAGGCAGCAAGTAA 120

4060 CTTCCCTTTCAGTACATGTATTTCCTGAAAGACCTGTGGCAAAGAGGCAGCAAGTAA 4119

121 AAGTAGATATCCTTGTGCTTCCATTTCAGCTATAAAATTGGAACCGTTAGAC 180

4120 AAGTAGATATCCTTGTGCTTTCCTTCGATTTCAGCTATAAAATTGGAACCGTTAGAC 4179

181 TGCCACGAGAAATGCCATGGTTGTGAGAAGATTAAACATTTCTGGGTTAGTGAATAGCATTCA 240

4180 TGCCACGAGAAATGCATGGTTGTGTGAGAAGATTAAACATTTCTGGTTAGTGAATAGCATTTCA 4239

241 TACGCTTTTGGGCACCTTCCCCCTGCAACTTGCCAGATAAGCACTATTTCAGCTCTTATTC 300

4240 TACGCTTTTGGGCACCTTCCCCCTGCAACTTGCCAGATAAGCACTATTTCAGCTCTTATTCC 4299

301 CAGTCTGACATCAGCAAGTGTGATTTTCTATGAAAAATTCTACTATGACTCCTTATTTA 360

4300 CAGTCTGACATCAGCAAGTGTGATTTTCTATGAAAAATTCTACTATGACTCCTTATTTA 4359

361 AGTATACAAGAACTTGTGACTCAGAAGATAATATTACAGAGTGGAACCCCTAG 420

QY	1501	GTGTGCATCTGCTACTGGCTTGAACAGATGGCAGAACACACAGATAGATGAGAGTTTC	1560	QY	2581	GACATTAATTTGCTGTTGTTAAATATATATGCTCTTTTGTGTTTCTGTCACCAAGGATGGAC	2640
Db	5500	GTGTGCATCTGCTACTGGCTTGAACAGATGGCAGAACACACAGATAGATGAGAGTTTC	5559	Db	6580	GACATTAATTTGCTGTTGTTAAATATATATGCTCTTTTGTGTTTCTGTCACCAAGGATGGAC	6639
QY	1561	TCCATTTTGTGTTGTTCTGGAGCTCATGTATATTCAGAGACACAAAGGTGGAGAGAGC	1620	QY	2641	AGTGATTCAGAACCGTCAAGACGGTAGTGTGACTTTTGGCAGGAATATGGATCCATATAA	2700
Db	5560	TCCATTTTGTGTTGTTCTGGAGCTCATGTATATTCAGAGACACAAAGGTGGAGAGAGC	5619	Db	6640	AGTGATTCAGAACCGTCAAGACGGTAGTGTGACTTTTGGCAGGAATATGGATCCATATAA	6699
QY	1621	TTTGTTCATCTCTTAGCAGATAAACGCTCTCAAACTGGGTGGACTTACTATAAGTAAA	1680	QY	2701	ACAGGATTTTGAAGATTTGCAACCAACACACATGGGAGAAATTTACTGTGGCTACACAG	2760
Db	5620	TTTGTTCATCTCTTAGCAGATAAACGCTCTCAAACTGGGTGGACTTACTATAAGTAAA	5679	Db	6700	ACAGGATTTTGAAGATTTGCAACCAACACACATGGGAGAAATTTACTGTGGCTACACAG	6759
QY	1681	ATGAAATCTAATATTTGTTATATTTTCAAGGCTTATATTAACACACTCCCTTAGTAA	1740	QY	2761	TAACGAACAGGCATGCAGAAATATAATCAATCTTATTTGAAATGGGATTTTTTTTAAATAA	2820
Db	5680	ATGAAATCTAATATTTGTTATATTTTCAAGGCTTATATTAACACACTCCCTTAGTAA	5739	Db	6760	TAACGAACAGGCATGCAGAAATATAATCAATCTTATTTGAAATGGGATTTTTTTTAAATAA	6819
QY	1741	ACTTATGTAATGTTATTTAAAGAAATGTGTACTAAATACAAAGTAATTTATGTCTAAAC	1800	QY	2821	AAACATTTCAATGTTTGGAGCCCTGTTTATAGGCAAGTTTAAGAGGAGTTTCTGACAAAATGT	2880
Db	5740	ACTTATGTAATGTTATTTAAAGAAATGTGTACTAAATACAAAGTAATTTATGTCTAAAC	5799	Db	6820	AAACATTTCAATGTTTGGAGCCCTGTTTATAGGCAAGTTTAAGAGGAGTTTCTGACAAAATGT	6879
QY	1801	CCCTGAACATAATGTTGCTTACATTTGACAGATGTGAGGAAATATCAGGAAAGAGGT	1860	QY	2881	GGAACTAAAGATAAGGGAAGAAAGCGAGTTTATAGTTTCCCAAAATTTTATTTTGGTG	2940
Db	5800	CCCTGAACATAATGTTGCTTACATTTGACAGATGTGAGGAAATATCAGGAAAGAGGT	5859	Db	6880	GGAACTAAAGATAAGGGAAGAAAGCGAGTTTATAGTTTCCCAAAATTTTATTTTGGTG	6939
QY	1861	GAACATCTGAAATGTAATCTCAATCAACCTGACAGTTCTGTCAAAACCGTATAGATATAC	1920	QY	2941	AGAGATTTTATTTGTTTTCTTTTATAGGTGAATATTTGGCTTGGAAATGATAAAATTAGCC	3000
Db	5860	GAACATCTGAAATGTAATCTCAATCAACCTGACAGTTCTGTCAAAACCGTATAGATATAC	5919	Db	6940	AGAGATTTTATTTGTTTTCTTTTATAGGTGAATATTTGGCTTGGAAATGATAAAATTAGCC	6999
QY	1921	TGTGACATGAATACAGAAATGGAGTAAGCTTTTCAGAGTTGTTGACCTGTTGATCTGT	1980	QY	3001	AGCTTACAGAGTGGAGCCACAGAACTTTTGCATAGAAATGGAGGACTGGAAAGGAGACA	3060
Db	5920	TGTGACATGAATACAGAAATGGAGTAAGCTTTTCAGAGTTGTTGACCTGTTGATCTGT	5979	Db	7000	AGCTTACAGAGTGGAGCCACAGAACTTTTGCATAGAAATGGAGGACTGGAAAGGAGACA	7059
QY	1981	AAATATTTGGAATACCGTAAATGCGCAGGAAACAGGCCAGGTGTGTGCTCATACCTGT	2040	QY	3061	AAATTAAGGCTCAGCTATGAGGATTTCACTGTACAGAAATGAAGCCCAACAAATACAGATCT	3120
Db	5980	AAATATTTGGAATACCGTAAATGCGCAGGAAACAGGCCAGGTGTGTGCTCATACCTGT	6039	Db	7060	AAATTAAGGCTCAGCTATGAGGATTTCACTGTACAGAAATGAAGCCCAACAAATACAGATCT	7119
QY	2041	AAATCCAGCACCTTGGAGGCCAAAGTGGGCTGATAGCTTGAGCCCTAGGAGTTTGAACCT	2100	QY	3121	CAGTGAACAAATACAGAGGAAACAGCCGGTAAATGCCCTCATGGATGGAGCATCTCAGCTGA	3180
Db	6040	AAATCCAGCACCTTGGAGGCCAAAGTGGGCTGATAGCTTGAGCCCTAGGAGTTTGAACCT	6099	Db	7120	CAGTGAACAAATACAGAGGAAACAGCCGGTAAATGCCCTCATGGATGGAGCATCTCAGCTGA	7179
QY	2101	AGCCTGGGCAACATATGAGACCCCTAACTCTCAAAAAAATAAAAAAATAACAAAAA	2160	QY	3181	TGGGAGAAACAGGACCATGACCATTCACACGGCATGTTCTTAGCAGGATGACAGAG	3240
Db	6100	AGCCTGGGCAACATATGAGACCCCTAACTCTCAAAAAAATAAAAAAATAACAAAAA	6159	Db	7180	TGGGAGAAACAGGACCATGACCATTCACACGGCATGTTCTTAGCAGGATGACAGAG	7239
QY	2161	AAAAAATCAGCTGTGTTGTTAGTATGTCCTGTAGTCCAGCTATCCAGGAGGCTGAGA	2220	QY	3241	ACAATGACGGCTGTGTGTGGGCACTCTTGTCTCTCTGCTTAAATAACACACTAATATC	3300
Db	6160	AAAAAATCAGCTGTGTTGTTAGTATGTCCTGTAGTCCAGCTATCCAGGAGGCTGAGA	6219	Db	7240	ACAATGACGGCTGTGTGTGGGCACTCTTGTCTCTCTGCTTAAATAACACACTAATATC	7299
QY	2221	TGGGAGATCACTGAGCCCAACCTGGAGTCTTGATCATGCTACTGAACTGTAGCCTGG	2280	QY	3301	ATTACTCAGAAATCAATTAACAATTTTTTAATAGCTACCACTTCCTGGGCACTTACTGTCA	3360
Db	6220	TGGGAGATCACTGAGCCCAACCTGGAGTCTTGATCATGCTACTGAACTGTAGCCTGG	6279	Db	7300	ATTACTCAGAAATCAATTAACAATTTTTTAATAGCTACCACTTCCTGGGCACTTACTGTCA	7359
QY	2281	GCAACAGAGGATGAGTCTGCTCTCAAAAAAATAAAAAAATAAAAAAGCCAGAAAC	2340	QY	3361	GCCAGTGTCTTAAGCTTTTATGCACTCTGAAAGCATTTCAACTATAAGGTAGACATT	3420
Db	6280	GCAACAGAGGATGAGTCTGCTCTCAAAAAAATAAAAAAATAAAAAAGCCAGAAAC	6339	Db	7360	GCCAGTGTCTTAAGCTTTTATGCACTCTGAAAGCATTTCAACTATAAGGTAGACATT	7419
QY	2341	AAGACTTAGCTCTAACATCTAACATAGCTGACAAAGGAGTAATTTGATGTGAAATCAAC	2400	QY	3421	CTTATTTCTCAATTTACAGATGAGATTTAGAGAGATTTACGTGATTTTGTCCAATGTACACA	3480
Db	6340	AAGACTTAGCTCTAACATCTAACATAGCTGACAAAGGAGTAATTTGATGTGAAATCAAC	6399	Db	7420	CTTATTTCTCAATTTACAGATGAGATTTAGAGAGATTTACGTGATTTTGTCCAATGTACACA	7479
QY	2401	CTGATATTTAAAGTTATAAATATCTATAATTCACAAATTTGGGGTAAGATAAGCACTT	2460	QY	3481	ACTACCCAGAGATAAACTAGAAATTTGAGCAGCTTACTTTCTGAATATAGCAATTTAG	3540
Db	6400	CTGATATTTAAAGTTATAAATATCTATAATTCACAAATTTGGGGTAAGATAAGCACTT	6459	Db	7480	ACTACCCAGAGATAAACTAGAAATTTGAGCAGCTTACTTTCTGAATATAGCAATTTAG	7539
QY	2461	CGAGTTTCCAGAGATTTACAGTTTACCTCTCATATTTATTTCTTATTTGTCATTTT	2520	QY	3541	ATMAATCCTATATCTCTATATTTCAAAGTGTGTGAAACCTTTCAATTTCAATTTCCAG	3600
Db	6460	CGAGTTTCCAGAGATTTACAGTTTACCTCTCATATTTATTTCTTATTTGTCATTTT	6519	Db	7540	ATMAATCCTATATCTCTATATTTCAAAGTGTGTGAAACCTTTCAATTTCAATTTCCAG	7599
QY	2521	TAGAGCACCATAATATACTAAATGGAATGGACAGGGATTCAGATATATTTTCAAAGT	2580	QY	3601	GGTCTCTGATACCTAAAGGTTGTAAGCTTATTTCCAGTATATAAGTAAACAAACACAGT	3660
Db	6520	TAGAGCACCATAATATACTAAATGGAATGGACAGGGATTCAGATATATTTTCAAAGT	6579	Db	7600	GGTCTCTGATACCTAAAGGTTGTAAGCTTATTTCCAGTATATAAGTAAACAAACACAGT	7659
				QY	3661	CCCTAGATGGATGGCCAAAGGCCAGCTTATCTCTTTCTTGTGTAGGGACAGGAG	3720

Db 7660 CCTAGATGATGGCACAAAGCCAGTAATCTCTCTTTCTGCTATAGGCGCAGGAG 7719
QY 3721 GTCCTTTGGTATAGTGAGCTATAGTATAGCACCCAAAGAAAGACTACTGTGCACA 3780
Db 7720 GTCCTTTGGTATAGTGAGCTCTATGTATAGCACCCAAAGAAAGACTACTGTGCACA 7779
QY 3781 CGAGTGTAGCAGTCTTTTATGGTATCTGCAAAAGCTAACTTGACCGGTAGTCTCTGT 3840
Db 7780 CGAGTGTAGCAGTCTTTTATGGTATCTGCAAAAGCTAACTTGACCCGTAAGTCTCTGT 7839
QY 3841 TTCTATAACGCCAAACACATTTTCTTTTCAGGTTTAAACATCCAGAGAAACAGTGT 3900
Db 7840 TTCTATAACGCCAAACACATTTTCTTTTCAGGTTTAAACATCCAGAGAAACAGTGT 7899
QY 3901 CTAAGAAGACGGTGTGATGTGTATTAATAGATGTATGAGCAATCCAAACGGCA 3960
Db 7900 CTAAGAAGACGGTGTGATGTGTATTAATAGATGTATGAGCAATCCAAACGGCA 7959
QY 3961 GATACTACTGGGTGGACAGTACACCTGGGACATGCGAAGCATGCGACAGATGATG 4020
Db 7960 GATACTACTGGGTGGACAGTACACCTGGGACATGCGAAGCATGCGACAGATGATG 8019
QY 4021 TAGTATGATGAATTTGGAAGGGTCACTGTTACTCTCAATGAGGAAGATGAGTATCAAGATCA 4080
Db 8020 TAGTATGATGAATTTGGAAGGGTCACTGTTACTCTCAATGAGGAAGATGAGTATCAAGATCA 8079
QY 4081 GGGCCCTCTTCCACAGCAATAGTCCCAATAGCTAGATTTTTCCTCTCTGTATGTGAC 4140
Db 8080 GGGCCCTCTTCCACAGCAATAGTCCCAATAGCTAGATTTTTCCTCTCTGTATGTGAC 8139
QY 4141 AACATTTTGTACATATGTTATTTGAAATTTCTTTTCATACATATATTCCTCTAAACCT 4200
Db 8140 AACATTTTGTACATATGTTATTTGAAATTTCTTTTCATACATATATTCCTCTAAACCT 8199
QY 4201 CTCAAGCAGACGTGAGTGTGACTTTTGGAAAAAGTATAGGATAAATTAATTTAAATAG 4260
Db 8200 CTCAAGCAGACGTGAGTGTGACTTTTGGAAAAAGTATAGGATAAATTAATTTAAATAG 8259
QY 4261 CACATGATTTCTTTTGTGTTTCTTCAATTTCTCTGCTACCCAAAGTAACAAAAGTAT 4320
Db 8260 CACATGATTTCTTTTGTGTTTCTTCAATTTCTCTGCTACCCAAAGTAACAAAAGTAT 8319
QY 4321 AGTTTGTACAGAGTGGTCTCATATTTTCAGTTCAGTGTGATGGAGAAATTTTCAAT 4380
Db 8320 AGTTTGTACAGAGTGGTCTCATATTTTCAGTTCAGTGTGATGGAGAAATTTTCAAT 8379
QY 4381 AAGGAAGAGGGTCTTTTATCCTTGTGTAAGAAAAACCATGACGGAAGGAAAACTGAT 4440
Db 8380 AAGGAAGAGGGTCTTTTATCCTTGTGTAAGAAAAACCATGACGGAAGGAAAACTGAT 8439
QY 4441 GTTTAAAGTCACATTTTAAACTATATTTATTTATGTAAGATCTGTCAAGAAACTTC 4500
Db 8440 GTTTAAAGTCACATTTTAAACTATATTTATTTATGTAAGATCTGTCAAGAAACTTC 8499
QY 4501 CAAAAGATTTATTAATTAACACAGACTCTGTGCAATAGTTAATGTTTCTGTTTGG 4560
Db 8500 CAAAAGATTTATTAATTAACACAGACTCTGTGCAATAGTTAATGTTTCTGTTTGG 8559
QY 4561 TAATCCACACATTTCAATGAGTGTAGGCTTTGCACTTGTAGGAAGGAGAGCGTTTCAAC 4620
Db 8560 TAATCCACACATTTCAATGAGTGTAGGCTTTGCACTTGTAGGAAGGAGAGCGTTTCAAC 8619
QY 4621 CTCAAATAGCTAATAAACCGGCTTGAATTTTGAAGATTTAAATCTGACTTAGGACG 4680
Db 8620 CTCAAATAGCTAATAAACCGGCTTGAATTTTGAAGATTTAAATCTGACTTAGGACG 8679
QY 4681 GGCACGGTGGCTCACGACTATATCCACACTTTGGGGGCTGAGCGCGGCTCACAA 4740
Db 8680 GGCACGGTGGCTCACGACTATATCCACACTTTGGGGGCTGAGCGCGGCTCACAA 8739
QY 4741 GGTTCAGGAGTTCAAGACACGAGCTGACCAATATGGTGAACCCCATCTCTACTAAAAATAC 4800

Db 8740 GGTCCAGGAGTTCAAGACAGAGCTGACCAATATGGTGAACCCCATCTCTACTAAAAATAC 8799
QY 4801 AAAAATTAGCCAGGGGTGGTGGCAGGTGCTTAGTCCAGCTAGCCTGTGAGGTGGAG 4860
Db 8800 AAAAATTAGCCAGGGGTGGTGGCAGGTGCTTAGTCCAGCTAGCCTGTGAGGTGGAG 8859
QY 4861 ATTGCATTGAGCCAAGATC 4879
Db 8860 ATTGCATTGAGCCAAGATC 8878
RESULT 6
ABX63814
ID ABX63814 standard; cDNA; 8674 BP.
XX AC ABX63814;
XX DT 26-FEB-2003 (first entry)
XX DE Human cDNA #814 differentially expressed in activated vascular tissue.
XX KW Human; gene; ss; vascular tissue; cytostatic; atherosclerosis; cardiant;
XX KW hypotensive; antidiabetic; gynaecological; vasotropic; cerebroprotective;
XX KW gene therapy; vascular disease; cancer; coronary; artery disease;
XX KW hypertension; diabetes; pre-eclampsia; restenosis;
XX KW ischaemia-reperfusion injury; stroke.
XX OS Homo sapiens.
XX PN US2002137081-A1.
XX PD 26-SEP-2002.
XX PF 08-JAN-2002; 2002US-00044090.
XX PR 28-JUL-2000; 2000US-0222469P.
XX PR 08-JAN-2001; 2001US-0260483P.
XX PA (BAND/) BANDMAN O.
XX PI Bandman O;
XX WPI; 2003-110597/10.
XX CC Combination for diagnosing, staging, treating, or monitoring the
XX CC progression of treatment of a vascular disease, e.g. atherosclerosis,
XX CC comprises several cDNAs that are differentially expressed in activated
XX CC vascular tissue.
XX PS Claim 1; Page; 18pp; English.
XX CC This invention relates to a combination comprising several cDNAs that are
XX CC differentially expressed in activated vascular tissue. The invention also
XX CC discloses a high throughput method for detecting differentially expressed
XX CC cDNAs in a sample. The cDNAs of the invention may have
XX CC antiarteriosclerotic; cytostatic; cardiant; hypotensive; antidiabetic;
XX CC gynaecological; vasotropic and cerebroprotective activities and may be
XX CC used in gene therapy. The cDNAs of the invention may be used in a high-
XX CC throughput methods for detecting differential expression of one or more
XX CC cDNAs in a sample, or screening several molecules or compounds to
XX CC identify a molecule or compound that specifically binds a cDNA of the
XX CC invention. A protein encoded by the cDNA may be used to screen several
XX CC molecules or compounds to identify a ligand that specifically binds to
XX CC the protein, or to produce or purify an antibody to the protein that can
XX CC be used to detect a protein in a sample or purify a natural or
XX CC recombinant protein from a sample. The nucleotides may be useful for
XX CC diagnosing, staging, treating, or monitoring the progression of treatment
XX CC of a vascular disease, e.g. atherosclerosis, cancer, coronary artery
XX CC disease, hypertension, diabetes, pre-eclampsia, ischaemia- reperfusion
XX CC injury, restenosis, or stroke. The cDNAs can also be used for large-scale
XX CC genetic or gene expression analysis of several new nucleic acid
XX CC molecules. Antibodies to the proteins encoded by the cDNAs are useful for
XX CC diagnosing pre-pathologic disorders, and chronic or acute diseases

CC associated with abnormalities in the expression, amount or distribution
CC of the protein. The present sequence represents a cDNA of the invention
CC that is differentially expressed in activated vascular tissue. Note: The
CC sequence data for this patent did not form part of the specification, but
CC was obtained in electronic format directly from USPTO at
CC <http://seqdata.uspto.gov/sequence.html?docID=20020137081>
XX

SQ Sequence 8674 BP; 2658 A; 1311 C; 1485 G; 2640 T; 0 U; 580 Other;

Query Match	87.5%;	Score 4268.2;	DB 8;	Length 8674;
Best Local Similarity	92.7%;	Pred No. 0;		
Matches 4336;	Conservative 0;	Mismatches 336;	Indels 6;	Gaps 6;

QY	1	TCAGAAATAGTGTTCATGAGTTAAATAACAATGTGGAAGCTGTTTCCAGACCTCTCTT	60
DB	4000	TCAGAAATAGTGTTCATGAGTTAAATAACAATGTGGAAGCTGTTTCCAGACCTCTCTT	4059
QY	61	CTTCCCTTTTCAGTACATGATTTGCTGGAAGACCTGTGCAAGAGGAGGAGCAAGTAA	120
DB	4060	CTTCCCTTTTCAGTACATGATTTGCTGGAAGACCTGTGCAAGAGGAGGAGCAAGTAA	4119
QY	121	AAGGTAGATATCTTGTGCTTCCATTCGATTTTCAGCTATAAAATTCGAACCGTTAGAC	180
DB	4120	AAGGTAGATATCTTGTGCTTCCATTCGATTTTCAGCTATAAAATTCGAACCGTTAGAC	4179
QY	181	TGCCACGAGAATGCTGTTGTGAGAGATTAACTATTTCTGGGTTAGTGAATGACATTC	240
DB	4180	TGCCACGAGAATGCTGTTGTGAGAGATTAACTATTTCTGGGTTAGTGAATGACATTC	4239
QY	241	TACGCTTTTGGGACCTTCCCTGCACTTGCAGATTAAGCACTTTCAGCTCTTATTC	300
DB	4240	TACGCTTTTGGGACCTTCCCTGCACTTGCAGATTAAGCACTTTCAGCTCTTATTC	4299
QY	301	CAGTCTGACATCAGCAAGTGTGATTTCTGAAATAATTTCTACTATGACTCTCTATTTTA	360
DB	4300	CAGTCTGACATCAGCAAGTGTGATTTCTGAAATAATTTCTACTATGACTCTCTATTTTA	4359
QY	361	AGTATACAAGAACTTGTGACTCAGAAGATAATATTTACAGAGTGGAAAAAACCCCTAG	420
DB	4360	AGTATACAAGAACTTGTGACTCAGAAGATAATATTTACAGAGTGGAAAAAACCCCTAG	4419
QY	421	CATTTATAGTTTAACTTTGAGTTTGAATGAGAGATTTATCCATATATATTTCAATT	480
DB	4420	CATTTATAGTTTAACTTTGAGTTTGAATGAGAGATTTATCCATATATATTTCAATT	4479
QY	481	GTGTTGTGGATAATGACACCTTAACCTGTGAATCTTGAAGTCAAGATTTGAGTGTGTTG	540
DB	4480	GTGTTGTGGATAATGACACCTTAACCTGTGAATCTTGAAGTCAAGATTTGAGTGTGTTG	4539
QY	541	ACTTGGTGGTCAGGAAAACAGCTAGTGTGAGCCTGGCACAGGCATCTCAGTGAAGCA	600
DB	4540	ACTTGGTGGTCAGGAAAACAGCTAGTGTGAGCCTGGCACAGGCATCTCAGTGAAGCA	4599
QY	601	TACCCACAGTTGGAAATTTTCAAGAAATCAAGGAATCATGACATCTTATAAATTTCA	660
DB	4600	TACCCACAGTTGGAAATTTTCAAGAAATCAAGGAATCATGACATCTTATAAATTTCA	4659
QY	661	AGGTTCTGCTATACCTATGTAATGGATAATTAATCAAGCATATCCACTCTGAAGAT	720
DB	4660	AGGTTCTGCTATACCTATGTAATGGATAATTAATCAAGCATATCCACTCTGAAGAT	4719
QY	721	TGAATCTCTCAGATGGAGAACCCCAATCTGCTTTCTCTCTTTTCCCTCACCAAGAAA	780
DB	4720	TGAATCTCTCAGATGGAGAACCCCAATCTGCTTTCTCTCTTTTCCCTCACCAAGAAA	4779
QY	781	TAAACAACCTATTTCAATTTTACTGGACACAACTTTTAGCGTATACCTATGTAATTTA	840
DB	4780	TAAACAACCTATTTCAATTTTACTGGACACAACTTTTAGCGTATACCTATGTAATTTA	4839
QY	841	CTAGTATGGTGGTTAGGATTTATGTAATTTGTATATGTCATGCGCCCAAAATCAATTC	900
DB	4840	CTAGTATGGTGGTTAGGATTTATGTAATTTGTATATGTCATGCGCCCAAAATCAATTC	4899

QY	901	TAAATATGACATATATATATCAATCACTGCTTGGTGATAGCTAGTGTGTTAATAGTTTATCTC	960
DB	4900	TAAATATGACATATATATATCAATCACTGCTTGGTGATAGCTAGTGTGTTAATAGTTTATCTC	4959
QY	961	AGAAATCAAAATTTGTATAGTTAAATACATTAAGTTTATGAGGCAAAATCTCACTAAT	1020
DB	4960	AGAAATCAAAATTTGTATAGTTAAATACATTAAGTTTATGAGGCAAAATCTCACTAAT	5019
QY	1021	TCTACATAATTTCAATTTTCCAGATAATGAAATAGTAGTCAA- TGAGTACTCTCTCAGAAC	1079
DB	5020	TCTACATAATTTCAATTTTCCAGATAATGAAATAGTAGTCAA- TGAGTACTCTCTCAGAAC	5079
QY	1080	TGGAAGACACCAATATATATATAGATGAGACTGTGAATAGCAATATCCCACTAACCTTC	1139
DB	5080	TGGAAGACACCAATATATATATAGATGAGACTGTGAATAGCAATATCCCACTAACCTTC	5139
QY	1140	GTGTGCTTTCGTTCAATCTCGAAACCTGGAAGCAAAATACAAAGTTAGAACTGATG	1199
DB	5140	GTGTGCTTTCGTTCAATCTCGAAACCTGGAAGCAAAATACAAAGTTAGAACTGATG	5199
QY	1200	TCTCAGCTCAAATGGAAATTTGTGCGACCCCATGCTGTGAGTTGCAATATTCCTGTGG	1259
DB	5200	TCTCAGCTCAAATGGAAATTTGTGCGCA-CCCATGCACTGTGAGTTGCAATATTCCTGTGG	5258
QY	1260	TGCTGGCAAAAGTTAACTGATTCATAAACAATATTTTATAGAGTTCCAGAGAACTCACA	1319
DB	5259	TGCTGGCAAAAGTTAACTGATTCATAAACAATATTTTATAGAGTTCCAGAGAACTCACA	5318
QY	1320	CACCAAAATTAAGAGAACCAACAACAACCAAAATGCTAAGTGATTTTCCCAACAGAT	1379
DB	5319	CACCAAAATTAAGAGAACCAACAACAACCAAAATGCTAAGTGATTTTCCCAACAGAT	5378
QY	1380	CATAATGACATTAACAGTACATATAAAATATCTTAGCCAGTTGTGTTTGGACTGGCC	1439
DB	5379	CATAATGACATTAACAGTACATATAAAATATCTTAGCCAGTTGTGTTTGGACTGGCC	5438
QY	1440	TGTTGCAATTTGCTGTTTGTATGAGCAGGATGGGGCAAGGTAGTCCAGGGGTGGCTGA	1499
DB	5439	TGTTGCAATTTGCTGTTTGTATGAGCAGGATGGGGCAAGGTAGTCCAGGGGTGGCTGA	5498
QY	1500	TGTTGCAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1559
DB	5499	TGTTGCAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	5558
QY	1560	CTCCATTTTGTGTTTCTGGAGCTCATGATATTTCCAGGACACAAAAGGTGGAGAAGAG	1619
DB	5559	CTCCATTTTGTGTTTCTGGAGCTCATGATATTTCCAGGACACAAAAGGTGGAGAAGAG	5618
QY	1620	CTTTGTTCACTCTCTTAGCAGATAAAGCTCTCAAACTGGGTGGAGTTTACTTAAGTAA	1679
DB	5619	CTTTGTTCACTCTCTTAGCAGATAAAGCTCTCAAACTGGGTGGAGTTTACTTAAGTAA	5678
QY	1680	AATGAAATCTAAATTTTGTATATTTTCAAGAGTCTATAATAACACACTCTCTAGT	1739
DB	5679	AATGAAATCTAAATTTTGTATATTTTCAAGAGTCTATAATAACACACTCTCTAGT	5738
QY	1740	AACCTATGTAATTTTAAAGAAATTTGTCATCTAAATACAAAGTAAATTTATGTCATAA	1799
DB	5739	AACCTATGTAATTTTAAAGAAATTTGTCATCTAAATACAAAGTAAATTTATGTCATAA	5798
QY	1800	CCCCTGAACTAAATGTTGTTCTTACATTTTGCAGAAATGTCAGGAAATATCAGGAAAGAG	1859
DB	5799	CCCCTGAACTAAATGTTGTTCTTACATTTTGCAGAAATGTCAGGAAATATCAGGAAAGAG	5858
QY	1860	TGAAACATCTGAAATGTTATCTCATTCACCTGACAGTTTCTGTCAAACCGTATAGAGTATA	1919
DB	5859	TGAAACATCTGAAATGTTATCTCATTCACCTGACAGTTTCTGTCAAACCGTATAGAGTATA	5918
QY	1920	CTGTGACATGAATACAGAAATGGAAGTTAGCTTTCGACAGTTGTGACCTGTTGATCTG	1979
DB	5919	CTGTGACATGAATACAGAAATGGAAGTTAGCTTTCGACAGTTGTGACCTGTTGATCTG	5978
QY	1980	TAAATTTTGGTACCGTAAATGCCAGAAACAGGCCAGGTGTGGTGTGCTCATACCTG	2039

QY 4198 ACTCTCAAGCAGACGCTGAGTGTGACTTTTGAAGAAAGTATAGGATAAATACATTAAAA 4257
Db 8198 ACTCTCAAGCAGACGCTGAGTGTGACTTTTGAAGAAAGTATAGGATAAATACATTAAAA 8257
QY 4258 TAGCACATGATTTCTTTTGTGTTCTTCTCATTTCTCTGCTACCAAGAGTAACAAAAG 4317
Db 8258 TAGCACATGATTTCTTTTGTGTTCTTCTCATTTCTCTGCTCA-CCAAGAAAGTAACAAAAG 8316
QY 4318 TATAGTCTTTCAGACAGTGTGCTGCTCATATTTCAATTTCACTTCTAGTTGATTCGCGAATTTCA 4377
Db 8317 TATAGTCTTTCAGACAGTGTGCTGCTCATATTTCAATTTCACTTCTAGTTGATTCGCGAATTTCA 8376
QY 4378 AATAAGGAAGGGGCTCTTTTATCTCTGCTAGGAAACCATGACGGAAGGAAAGAAACT 4437
Db 8377 AATAAGGAAGGGGCTCTTTTATCTCTGCTAGGAAACCATGACGGAAGGAAAGAAACT 8436
QY 4438 GATGTTTAAAGTCCACTTTTAAACTATATTTATTTATGATGATCTGTCTAAAGAAAC 4497
Db 8437 GATGTTTAAAGTCCACTTTTAAACTATATTTATTTATGATGATCTGTCTAAAGAAAC 8496
QY 4498 TTCCAAAAGATTTTATTAATTAACACGAGCTCTGTTGCAATAAGTTAATGTTTCTTCTGT 4557
Db 8497 TTCCAAAAGATTTTATTAATTAACACGAGCTCTGTTGCAATAAGTTAATGTTTCTTCTGT 8556
QY 4558 TTGTAATCCACACATTCATGAGTTAGGCTTTTGCACTTTGTAAGGAAGGAGCGTTTCA 4617
Db 8557 TTGTAATCCACACATTCATGAGTTAGGCTTTTGCACTTTGTAAGGAAGGAGCGTTTCA 8616
QY 4618 AACCTCAATAGCTAATAACCGGCTTGAATATTTGAAGATTTAAATCTGACTCTA 4675
Db 8617 AACCTCAATAGCTAATAACCGGCTTGAATATTTGAAGATTTAAATCTGACTCTA 8674

RESULT 7

ABX91986
ID ABX91986 standard; cDNA; 4043 BP.
XX AC ABX91986;
XX DT 08-MAY-2003 (first entry)
XX DE Lung specific nucleic acid (LSNA) #28.
XX XX Human; gene therapy; vaccine; lung specific antigen; cancer diagnosis;
KW cancer monitoring; cancer staging; cancer imaging; lung cancer;
XX non-cancerous diseases of the lung; transgenic animal; gene; ss.
XX OS Homo sapiens.
XX PN WO200268633-A2.
XX PD 06-SEP-2002.
XX PF 21-NOV-2001; 2001WO-US043612.
XX PR 22-NOV-2000; 2000US-0252500P.
XX PA (DIAD-) DIADEXUS INC.
XX PI Macina RA, Recipon H, Chen S, Sun Y, Liu C;
XX DR WPI; 2002-713376/77.
XX PT New isolated human nucleic acid molecule and polypeptide, useful for
PT identifying, diagnosing, monitoring, staging, imaging and treating lung
PT cancer and non-cancerous diseases of the lung.
XX PS Claim 1; Page 181-183; 389pp; English.
XX XX The invention describes an isolated human nucleic acid (I) encoding any
CC of 120-1533 residue amino acid sequences (SI), given in the
CC specification, comprising any of 164 179-12421 base pair sequences (S2),

CC given in the specification. The methods and compositions of the present
CC invention are useful for identifying, diagnosing, monitoring, staging,
CC imaging and treating lung cancer and non-cancerous diseases of the lung.
CC They are also used for identifying lung tissue, monitoring and
CC identifying and/or designing antagonists of the polypeptide of the
CC invention, gene therapy, production of transgenic animals and production
CC of engineered lung tissue for treatment and research. This sequence
CC encodes a lung specific nucleic acid
XX SQ Sequence 4043 BP; 1283 A; 824 C; 842 G; 1094 T; 0 U; 0 Other;
Query Match 42.7%; Score 2081.2; DB 6; Length 4043;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2083; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 2018 CAGGTGTGCTGCTCATACCTGTAAATCCACAGCCTTGGGAGGCCAAAGTGGCTGATAG 2077
Db 1958 CAAGTGTGCTGCTCATACCTGTAAATCCACAGCCTTGGGAGGCCAAAGTGGCTGATAG 2017
QY 2078 CTTGAGCCTTAGGAGTTTGAAACTAGCCTGGGCAACATAATGAGACCCCTAACTCTACAAA 2137
Db 2018 CTTGAGCCTTAGGAGTTTGAAACTAGCCTGGGCAACATAATGAGACCCCTAACTCTACAAA 2077
QY 2138 AAAAAAATACCAAAAAAATAAATCAGCTGTGTGGTAGTATGCTGTGTAGT 2197
Db 2078 AAAAAAATACCAAAAAAATAAATCAGCTGTGTGGTAGTATGCTGTGTAGT 2137
QY 2198 CCCAGCTATCCAGAGGCTGAGATGGGAGATCAGCTGAGCCCAACCTGGAGTCTTCAT 2257
Db 2138 CCCAGCTATCCAGAGGCTGAGATGGGAGATCAGCTGAGCCCAACCTGGAGTCTTCAT 2197
QY 2258 CATGCTACTGAACCTGTAGCCTGGGCAACAGAGGATAGTGTAGTCTCTCAAAAAA 2317
Db 2198 CATGCTACTGAACCTGTAGCCTGGGCAACAGAGGATAGTGTAGTCTCTCAAAAAA 2257
QY 2318 AATTAATTAAGCCAGGAAACAAAGCTTAGCTCTTAACATCTAACATAGCTGACAAAG 2377
Db 2258 AATTAATTAAGCCAGGAAACAAAGCTTAGCTCTTAACATCTAACATAGCTGACAAAG 2317
QY 2378 AGTAATTTGATGTGGAATTCACCTGATATTTAAAGTTATAAATATCTAATTTCA 2437
Db 2318 AGTAATTTGATGTGGAATTCACCTGATATTTAAAGTTATAAATATCTAATTTCA 2377
QY 2438 ATTTGGGTAAGATAAAGCACTTCGAGTTTCCAAAGATTTTCAAGTTTACCTCTCATAT 2497
Db 2378 ATTTGGGTAAGATAAAGCACTTCGAGTTTCCAAAGATTTTCAAGTTTACCTCTCATAT 2437
QY 2498 TTATTTCTTATTTGTGCTATTTTAGAGCAACCAATATATATCTAAATGGAATGGACAGG 2557
Db 2438 TTATTTCTTATTTGTGCTATTTTAGAGCAACCAATATATATCTAAATGGAATGGACAGG 2497
QY 2558 GATTCAGATATTTTTCAAAGTGCATTTATTTCTGTTGGTTAATATCTCTTTTGG 2617
Db 2498 GATTCAGATATTTTTCAAAGTGCATTTATTTCTGTTGGTTAATATCTCTTTTGG 2557
QY 2618 TTTCTGTCAACCAAGGATGGACAGTGTATTTCAGAACCCGTCAAGCGTAGTGTGACTTT 2677
Db 2558 TTTCTGTCAACCAAGGATGGACAGTGTATTTCAGAACCCGTCAAGCGTAGTGTGACTTT 2617
QY 2678 GGCAGGAATGGGATCCATATAAACAGGATTTGGAAATTTGGCAACCAACACAGATGGG 2737
Db 2618 GGCAGGAATGGGATCCATATAAACAGGATTTGGAAATTTGGCAACCAACACAGATGGG 2677
QY 2738 AAGAAATTTACTGTGCTTACCAAGTAAACGACAGCATGCAAAATATAATCTATTATTG 2797
Db 2678 AAGAAATTTACTGTGCTTACCAAGTAAACGACAGCATGCAAAATATAATCTATTATTG 2737
QY 2798 AAATGGGATTTTTTTTAAATTAATAAACAATTTGGAAGCCTGTTTTAGGAGTTAA 2857
Db 2738 AAATGGGATTTTTTTTAAATTAATAAACAATTTGGAAGCCTGTTTTAGGAGTTAA 2797
QY 2858 GAGAGTTTCTGTGCAAAAAATGCTGAAGCTTAAAGATTAAGGAAGAAAGCGAGTTTACT 2917

Db 2428 CTGTTAAACATCAGATCCAGAAAAACAGTGTCTAAAGAACGCGTGTGGATGGTA 2369
QY 3929 TAATAGATGTCATGACAGCAATCCAAACGGCAGATACACTGGGTGGACAGTACACCTG 3988
Db 2368 TAATAGATGTCATGACAGCAATCCAAACGGCAGATACACTGGGTGGACAGTACACCTG 2309
QY 3989 GGACATGGCAAAAGCATGGCACAGATGATGCTGTAGTATGATGATGAATTTGGAAGGGTCAATG 4048
Db 2308 GGACATGGCAAAAGCATGGCACAGATGATGCTGTAGTATGATGATGAATTTGGAAGGGTCAATG 2249
QY 4049 GTACTCAATAGGAAGATGAGTATGAAGATCAGGCGCTTCTCCACAGCAATAGTCCCC 4108
Db 2248 GTACTCAATAGGAAGATGAGTATGAAGATCAGGCGCTTCTCCACAGCAATAGTCCCC 2189
QY 4109 AATAGCTAGATTTTCTCTCTGTATGTGACAAATTTTGTACATATGTTATTTGAA 4168
Db 2188 AATAGCTAGATTTTCTCTCTGTATGTGACAAATTTTGTACATATGTTATTTGAA 2129
QY 4169 TTTTCTTTTATATATATTTCTCTTAAACTCTCAAGCAGACGCTGAGTGTGACTTTTGTG 4228
Db 2128 TTTTCTTTTATATATATTTCTCTTAAACTCTCAAGCAGACGCTGAGTGTGACTTTTGTG 2069
QY 4229 AAAAAAGTATAGGATAAATTAATTAATTAATAGCAGATGTTTCTTTTGTCTTCTCAT 4288
Db 2068 AAAAAAGTATAGGATAAATTAATTAATTAATAGCAGATGTTTCTTTTGTCTTCTCAT 2009
QY 4289 TCTCTTGTCTACCCAGGAAGTAAACAAAGTATAGTTTTCACAGAGTGGTGTTCATAAT 4348
Db 2008 TCTCTTGTCTCA-CCAGGAAGTAAACAAAGTATAGTTTTCACAGAGTGGTGTTCATAAT 1950
QY 4349 TCAGTTCTAGTTGATTTGCGAGAAATTTTCAAAATAGGAAGAGGGGTCTTTTATCCTTTGCG 4408
Db 1949 TCAGTTCTAGTTGATTTGCGAGAAATTTTCAAAATAGGAAGAGGGGTCTTTTATCCTTTGCG 1890
QY 4409 TAGGAAACCATGACGGAAGGAAACCTGATGTTTAAAGTCCACTTTTAAACATATAT 4468
Db 1889 TAGGAAACCATGACGGAAGGAAACCTGATGTTTAAAGTCCACTTTTAAACATATAT 1830
QY 4469 TTATTTATGATGATCTGTCGAAGAAACCTTCAAAAGAGATTTTATTAATTAACAGACT 4528
Db 1829 TTATTTATGATGATCTGTCGAAGAAACCTTCAAAAGAGATTTTATTAATTAACAGACT 1770
QY 4529 CTGTTTCAATAGTTTAAATGTTTCTTTGTTTGTATTCACACATTCATGAGTTAGGCTT 4588
Db 1769 CTGTTTCAATAGTTTAAATGTTTCTTTGTTTGTATTCACACATTCATGAGTTAGGCTT 1710
QY 4589 TCACCTTTGGAAGGAGGAGGCTTCAAACTCAATAGCTAATAAACCGGTCTTGA 4648
Db 1709 TCACCTTTGGAAGGAGGAGGCTTCAAACTCAATAGCTAATAAACCGGTCTTGA 1650
QY 4649 TATTTGAAGATTTAAATCTGACTCTAGGACGGGACGCTGCTCAGCACTATAATCCCA 4708
Db 1649 TATTTGAAGATTTAAATCTGACTCTAGGACGGGACGCTGCTCAGCACTATAATCCCA 1590
QY 4709 ACATTTGGAGGCTGAGCGCGGCTCAAGAGTCAAGAGTTCAGACAGCTGACCA 4768
Db 1589 ACATTTGGAGGCTGAGCGCGGCTCAAGAGTTCAGACAGCTGACCA 1530
QY 4769 ATATGTTGAACCCCATCTCTACTAAATATCAAAATTAGCCGCGTGGGAGG 4828
Db 1529 ATATGTTGAACCCCATCTCTACTAAATATCAAAATTAGCCGCGTGGGAGG 1470
QY 4829 CTTGTAGGTCACAGCTAGCTGTGAGTGGAGATTCATTTAGCCAAAGATC 4879
Db 1469 CTTGTGA-GTCCACAGCTA-CTGTGAGTGGAGATTCATTTAGCCAAAGATC 1421

RESULT 9
ADL63650/c
ID ADL63650 standard; DNA; 4967 BP.
XX
AC ADL63650;

XX 20-MAY-2004 (first entry)
XX Human ovarian cancer DNA marker #21862.
XX Human; ovarian cancer; ds; tumour; cytostatic; DNA marker.
XX Homo sapiens.
XX WO200170979-A2.
XX 27-SEP-2001.
XX 21-MAR-2001; 2001WO-US009126.
XX 21-MAR-2000; 2000US-0191031P.
XX 25-MAY-2000; 2000US-0207124P.
XX 15-JUN-2000; 2000US-0211940P.
XX 07-JUL-2000; 2000US-0216820P.
XX 25-JUL-2000; 2000US-0220661P.
XX 21-DEC-2000; 2000US-0257672P.
XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX Lee J, Lillie J;
XX WPI; 2001-611502/70.
XX Novel isolated nucleic acid molecules (markers) overexpressed in ovarian cancer cells as compared to their normal non-cancerous ovarian cells are used to characterize stage, grade, histological type of ovarian cancer.
XX Disclosure; SEQ ID NO 21862; 106pp; English.
XX The invention relates to nucleic acid markers which are overexpressed in ovarian cancer cells as compared to their expression in normal (i.e. non-cancerous) ovarian cells. The invention also relates to polypeptides encoded by the markers, antibodies that selectively bind to the polypeptides, a method of inhibiting ovarian cancer in a patient at risk of developing ovarian cancer involving inhibiting expression of a gene corresponding to a marker of the invention and a method of treating a patient afflicted with ovarian cancer comprising providing to cells of the patient an antisense oligonucleotide complementary to a marker of the invention. The markers are useful for assessing if a patient is afflicted with ovarian cancer, which involves comparing the level of expression of a marker in a patient sample and a normal level of expression of the marker in a control non-ovarian cancer sample. A difference between the expression levels indicates ovarian cancer. The level of expression of a marker corresponds to a secreted protein or to a transcribed polynucleotide or its portion. The level of expression of the marker is assessed by detecting the presence in the sample, a protein or protein fragment corresponding to the marker. The presence of protein or protein fragment is detected using an antibody that specifically binds with the protein or protein fragment. Alternatively, the level of expression of the marker is assessed by detecting the presence of a transcribed polynucleotide which anneals with the marker or anneals with a portion of the polynucleotide comprising the marker, under stringent conditions. The marker is also used for monitoring the progression of ovarian cancer in a patient which involves detecting expression of the marker in a patient sample at a first point in time, repeating the method at a subsequent time and comparing the level of expression. The method is carried out using an ovarian tissue sample. A composition comprising a marker, polypeptide or antibody of the invention is used to treat ovarian cancer. This sequence represents a human ovarian cancer DNA marker of the invention.
XX Sequence 4967 BP; 1232 A; 1163 C; 1177 G; 1395 T; 0 U; 0 Other;
QY 3869 CAGGTTAATCATCAGATCCAGAAACAGTGTCTAAAGAGACGCGTGTGGATGGTA 3928

Query Match 19.9%; Score 971.8; DB 5; Length 4967;
Best Local Similarity 99.5%; Pred. No. 9.8e-194;
Matches 1006; Conservative 0; Mismatches 2; Indels 3; Gaps 3;
QY 3869 CAGGTTAATCATCAGATCCAGAAACAGTGTCTAAAGAGACGCGTGTGGATGGTA 3928

QY 4529 CTGTTGCAATAAGTTAATGTTTCTT 4554
 Db |||||
 1915 CTTGTCATAAATAATAGTTTTT 1940
 |||||
 RESULT 12
 AAK52610
 ID AAK52610 standard; cDNA; 2553 BP.
 XX
 AC AAK52610;
 XX
 DT 06-NOV-2001 (first entry)
 XX
 DE Human polynucleotide SEQ ID NO 2139.
 XX
 XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;
 KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
 KW tissue growth factor; immunomodulatory; cancer; leukaemia;
 KW nervous system disorder; arthritis; inflammation; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200157190-A2.
 XX
 PD 09-AUG-2001.
 XX
 PF 05-FEB-2001; 2001WO-US004098.
 XX
 PR 03-FEB-2000; 2000US-00496914.
 XX
 PR 27-APR-2000; 2000US-00560875.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 XX Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y;
 PI Ma Y, Zhao Q, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;
 PI Xue AJ, Yang Y, Wejhrman T, Goodrich R;
 XX
 DR WPI: 2001-476283/51.
 DR P-PSDB; AM79477.
 XX
 PT Nucleic acids encoding polypeptides with cytokine-like activities, useful
 in diagnosis and gene therapy.
 PS
 PS Claim 1; Page 4509-4510; 6221pp; English.
 XX
 CC The invention relates to polynucleotides (AAK51456-AAK53435) and the
 CC encoded polypeptides (AAW78323-AAW80302) that exhibit activity elating to
 CC cytokine, cell proliferation or cell differentiation or which may induce
 CC production of other cytokines in other cell populations. The
 CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
 CC peptide therapy. The polypeptides have various cytokine-like activities,
 CC e.g. stem cell growth factor activity, haematopoiesis regulating
 CC activity, tissue growth factor activity, immunomodulatory activity and
 CC activin/inhibin activity and may be useful in the diagnosis and/or
 CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
 CC inflammation. Note: Records for SEQ ID NO 2110 (AAK52581), 2111
 CC (AAK52582) and 3666 (AAW80020) are omitted as the relevant pages from the
 CC sequence listing were missing at the time of publication
 XX
 SQ Sequence 2553 BP; 812 A; 506 C; 546 G; 689 T; 0 U; 0 Other;
 Query Match 13.5%; Score 659.6; DB 4; Length 2553;
 Best Local Similarity 98.5%; Pred. No. 2.6e-128;
 Matches 676; Conservative 0; Mismatches 9; Indels 1; Gaps 1;
 QY 3869 CAGGTTAATACATCAGATCCAGAAAAACAGTGTCTAAAGAACGGTGGTGGTGGTA 3928

Db 1256 CTGGTTAATACATCAGATCCAGAAAAACAGTGTCTAAAGAACGGTGGTGGTGGTA 1315
 QY 3929 TAAATAGATGTCTATGAGCAATCCAAACGGCAGATACCTACCTGGGTGGACAGTACACCTG 3988
 Db 1316 TAAATAGATGTCTATGAGCAATCCAAACGGCAGATACCTACCTGGGTGGACAGTACACCTG 1375
 QY 3989 GGACATGGCAAGCATGGCAGATGATGGTGTAGTATGGATGAATGGAAAGGGTCTATG 4048
 Db 1376 GGACATGGCAAGCATGGCAGATGATGGTGTAGTATGGATGAATGGAAAGGGTCTATG 1435
 QY 4049 GTACTCAATGAGGAAGATGAGTATGAAGATCAGGCCCTCTTCCACAGCAATAGTCCCC 4108
 Db 1436 GTACTCAATGAGGAAGATGAGTATGAAGATCAGGCCCTCTTCCACAGCAATAGTCCCC 1495
 QY 4109 AATAGTATGATTTTCTCTCTCTATGTGACAAATTTTGTATCATTTATTTGGA 4168
 Db 1496 AATAGTATGATTTTCTCTCTCTATGTGACAAATTTTGTATCATTTATTTGGA 1555
 QY 4169 TTTTCTTCTATACATATATCT 4228
 Db 1556 TTTTCTTCTATACATATATCT 1615
 QY 4229 AAAAAAGTATAGGATAAATTTACATTTAAATAGCACATGATTTCTTTTCTTTCTTTCTT 4288
 Db 1616 AAAAAAGTATAGGATAAATTTACATTTAAATAGCACATGATTTCTTTTCTTTCTTTCTT 1675
 QY 4289 TCTCTTGTCTACCCCAAGAGTAACAAAAAGTATAGTATTTTGTACAGAGTGGTGTCTATAAT 4348
 Db 1676 TCTCTTGTCTCA - CCAAGAGTAACAAAAAGTATAGTATTTTGTACAGAGTGGTGTCTATAAT 1734
 QY 4349 TCAGTCTCTAGTGTGATGGAGAAATTTTCAATTAAGAGAGGGGTCTTTTATCCTTGTCTG 4408
 Db 1735 TCAGTCTCTAGTGTGATGGAGAAATTTTCAATTAAGAGAGGGGTCTTTTATCCTTGTCTG 1794
 QY 4409 TAGGAAAACCATGACGGAAAGGAAAACTGATGTTTAAAGTCCACTTTTAAACACTATAT 4468
 Db 1795 TAGGAAAACCATGACGGAAAGGAAAACTGATGTTTAAAGTCCACTTTTAAACACTATAT 1854
 QY 4469 TTATTTATGTAGGATCTGTCAAAGAAAACTTCCAAAAAGATTTATTAATAAACAGACT 4528
 Db 1855 TTATTTATGTAGGATCTGTCAAAGAAAACTTCCAAAAAGATTTATTAATAAACAGACT 1914
 QY 4529 CTGTTGCAATAAGTTAATGTTTCTT 4554
 Db |||||
 1915 CTTGTCATAAATAATAGTTTTT 1940
 |||||
 RESULT 13
 AAK51626
 ID AAK51626 standard; cDNA; 1862 BP.
 XX
 AC AAK51626;
 XX
 DT 06-NOV-2001 (first entry)
 XX
 DE Human polynucleotide SEQ ID NO 171.
 XX
 KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
 KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
 KW tissue growth factor; immunomodulatory; cancer; leukaemia;
 KW nervous system disorder; arthritis; inflammation; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200157190-A2.
 XX
 PD 09-AUG-2001.
 XX
 PF 05-FEB-2001; 2001WO-US004098.
 XX
 PR 03-FEB-2000; 2000US-00496914.
 PR 27-APR-2000; 2000US-00560875.

PR 20-JUN-2000; 2000US-00598075.
 PR 19-JUL-2000; 2000US-00620325.
 PR 01-SEP-2000; 2000US-00654936.
 PR 15-SEP-2000; 2000US-00663561.
 PR 20-OCT-2000; 2000US-00693325.
 PR 30-NOV-2000; 2000US-00728422.
 XX (HYSE-) HYSEQ INC.
 XX
 XX Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y;
 PI Ma Y, Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;
 PI Xue AJ, Yang Y, Wejhrman T, Goodrich R;
 XX
 DR WPI; 2001-476283/51.
 DR P-PSDB; AAM78493.
 XX
 XX Nucleic acids encoding polypeptides with cytokine-like activities, useful
 PT in diagnosis and gene therapy.
 XX
 XX Claim 1; Page 909-911; 6221pp; English.
 XX
 XX The invention relates to polynucleotides (AAK51456-AAK53435) and the
 CC encoded polypeptides (AAM78323-AAK80302) that exhibit activity elating to
 CC cytokine, cell proliferation or cell differentiation or which may induce
 CC production of other cytokines in other cell populations. The
 CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
 CC peptide therapy. The polypeptides have various cytokine-like activities,
 CC e.g. stem cell growth factor activity, haematopoiesis regulating
 CC activity, tissue growth factor activity, immunomodulatory activity and
 CC activin/inhibin activity and may be useful in the diagnosis and/or
 CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
 CC inflammation. Note: Records for SEQ ID NO 2110 (AAK52581), 2111
 CC (AAK52582) and 3666 (AAM80020) are omitted as the relevant pages from the
 CC sequence listing were missing at the time of publication
 XX
 XX Sequence 1862 BP; 602 A; 352 C; 426 G; 482 T; 0 U; 0 Other;
 SQ

Query Match 13.5%; Score 658.6; DB 4; Length 1862;
 Best Local Similarity 99.3%; Pred. No. 3.9e-128;
 Matches 672; Conservative 0; Mismatches 4; Indels 1; Gaps 1;

QY 3869 CAGGTTAATCATGATCCAGAAACAGTGTCTTAAGAAGACGGTGTGATGTGTA 3928
 DB 1182 CTGGTTAATCATGATCCAGAAACAGTGTCTTAAGAAGACGGTGTGATGTGTA 1241
 QY 3929 TAATAGATGTCTAGCAGCAATCCAAACGGCAGATACCTACTGGGTGGACAGTACACCTG 3988
 DB 1242 TAATAGATGTCTAGCAGCAATCCAAACGGCAGATACCTACTGGGTGGACAGTACACCTG 1301
 QY 3989 GGACATGGCAAGCATGGCAGATGATGGTGTAGTATGGATGAATCGAAGGGGTGATG 4048
 DB 1302 GGACATGGCAAGCATGGCAGATGATGGTGTAGTATGGATGAATCGAAGGGGTGATG 1361
 QY 4049 GTACTCAATGAGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 4108
 DB 1362 GTACTCAATGAGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1421
 QY 4109 AATAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 4168
 DB 1422 AATAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1481
 QY 4169 TTTTCTTTTAT 4228
 DB 1482 TTTTCTTTTAT 1541
 QY 4229 AAAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 4288
 DB 1542 AAAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1601
 QY 4289 TCTCTGTCTCAGCCAGAGTAAACAAAAGTATAGTATGATGATGATGATGATGATGATG 4348
 DB 1602 TCTCTGTCTCAGCCAGAGTAAACAAAAGTATAGTATGATGATGATGATGATGATGATG 1660

QY 4349 TCAGTTCTAGTTGATGCGAGAAATTTCAAAATAGGAAGAGGGGTCTTTTATCTTTGTCG 4408
 DB 1661 TCAGTTCTAGTTGATGCGAGAAATTTCAAAATAGGAAGAGGGGTCTTTTATCTTTGTCG 1720
 QY 4409 TAGGAACACCATGACGGAAGGAAAGAACTGATGTTTAAAGTCCACTTTTAAACTATAT 4468
 DB 1721 TAGGAACACCATGACGGAAGGAAAGAACTGATGTTTAAAGTCCACTTTTAAACTATAT 1780
 QY 4469 TTATTTATGATAGTCTGTCAAGAAAGAACTTCCAAAAGAGATTATTAATTAACCCAGACT 4528
 DB 1781 TTATTTATGATAGTCTGTCAAGAAAGAACTTCCAAAAGAGATTATTAATTAACCCAGACT 1840
 QY 4529 CTGTTGCAATAAGTTAA 4545
 DB 1841 CTGTTGCAATAAAAAA 1857

RESULT 14
 AAK51625
 ID AAK51625 standard; cDNA; 1976 BP.
 XX AC AAK51625;
 XX
 XX 06-NOV-2001 (first entry)
 XX
 XX Human polynucleotide SEQ ID NO 170.
 XX
 XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;
 KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
 KW tissue growth factor; immunomodulatory; cancer; leukaemia;
 KW nervous system disorder; arthritis; inflammation; ss.
 XX
 XX Homo sapiens.
 XX
 XX WO200157190-A2.
 XX
 XX 09-AUG-2001.
 XX
 XX 05-FEB-2001; 2001WO-US004098.
 XX
 XX 03-FEB-2000; 2000US-00496914.
 XX 27-APR-2000; 2000US-00560875.
 XX 20-JUN-2000; 2000US-00598075.
 XX 19-JUL-2000; 2000US-00620325.
 XX 01-SEP-2000; 2000US-00654936.
 XX 15-SEP-2000; 2000US-00663561.
 XX 20-OCT-2000; 2000US-00693325.
 XX 30-NOV-2000; 2000US-00728422.
 XX (HYSE-) HYSEQ INC.
 XX
 XX Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y;
 PI Ma Y, Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;
 PI Xue AJ, Yang Y, Wejhrman T, Goodrich R;
 XX
 XX WPI; 2001-476283/51.
 XX P-PSDB; AAM78492.
 XX
 XX Nucleic acids encoding polypeptides with cytokine-like activities, useful
 PT in diagnosis and gene therapy.
 XX
 XX Claim 1; Page 906-909; 6221pp; English.
 XX
 XX The invention relates to polynucleotides (AAK51456-AAK53435) and the
 CC encoded polypeptides (AAM78323-AAK80302) that exhibit activity elating to
 CC cytokine, cell proliferation or cell differentiation or which may induce
 CC production of other cytokines in other cell populations. The
 CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
 CC peptide therapy. The polypeptides have various cytokine-like activities,
 CC e.g. stem cell growth factor activity, haematopoiesis regulating
 CC activity, tissue growth factor activity, immunomodulatory activity and
 CC activin/inhibin activity and may be useful in the diagnosis and/or
 CC treatment of cancer, leukaemia, nervous system disorders, arthritis and

CC inflammation. Note: Records for SEQ ID NO 2110 (AAK52581), 2111
CC (AAK52582) and 3666 (AAK80020) are omitted as the relevant pages from the
CC sequence listing were missing at the time of publication
XX
SQ Sequence 1976 BP; 646 A; 368 C; 452 G; 510 T; 0 U; 0 Other;
Query Match 13.5%; Score 658.6; DB 4; Length 1976;
Best Local Similarity 99.3%; Pred. No. 3.9e-128;
Matches 672; Conservative 0; Mismatches 4; Indels 1; Gaps 1;
QY 3869 CAGGTTAAATCAGATCCAGAGAAACAGTGTCTTAAAGAGAGCGGTGGTGGTGA 3928
Db 1296 CTGGTTAAATCAGATCCAGAGAAACAGTGTCTTAAAGAGAGCGGTGGTGGTGA 1355
QY 3929 TAATAGATGTCTAGCAGCAATCCAAACGCGAGATCTACTGGGGTGGACAGTACACCTG 3988
Db 1356 TAATAGATGTCTAGCAGCAATCCAAACGCGAGATCTACTGGGGTGGACAGTACACCTG 1415
QY 3989 GGACATGGCAAGCATGGCACAGATGATGGTGTAGTATGGATGAATTTGGAAGGGTCTATG 4048
Db 1416 GGACATGGCAAGCATGGCACAGATGATGGTGTAGTATGGATGAATTTGGAAGGGTCTATG 1475
QY 4049 GTACTCAATGAGGAAGATGATGAAGATCAGGCCCTTCTTCCACAGCAATAGTCCCC 4108
Db 1476 GTACTCAATGAGGAAGATGATGAAGATCAGGCCCTTCTTCCACAGCAATAGTCCCC 1535
QY 4109 AATACGTAGATTTTGTCTCTCTGTATGTGACAAATTTTGTACATATGTTATTGGAA 4168
Db 1536 AATACGTAGATTTTGTCTCTCTGTATGTGACAAATTTTGTACATATGTTATTGGAA 1595
QY 4169 TTTTCTTTCATATPATATTCCTCTTAAACTCTCAAGCAGAGCGTGAAGTGTGACTTTTGG 4228
Db 1596 TTTTCTTTCATATPATATTCCTCTTAAACTCTCAAGCAGAGCGTGAAGTGTGACTTTTGG 1655
QY 4229 AAAAAGATGATGATTAATACATTAATTAATAGACATGATTTCTTTTGTGTTTCTTCATT 4288
Db 1656 AAAAAGATGATGATTAATACATTAATTAATAGACATGATTTCTTTTGTGTTTCTTCATT 1715
QY 4289 TCTCTGTCTCACCAGAGATGACAAAGTATAGTTTTCACAGAGTGGTGTGTTCAAT 4348
Db 1716 TCTCTGTCTCACCAGAGATGACAAAGTATAGTTTTCACAGAGTGGTGTGTTCAAT 1774
QY 4349 TCAGTTTCTAGTTGATTCGAGAAATTTTCAATAAGGAGAGGGTCTTTTATCCTTGTGCG 4408
Db 1775 TCAGTTTCTAGTTGATTCGAGAAATTTTCAATAAGGAGAGGGTCTTTTATCCTTGTGCG 1834
QY 4409 TAGGAACCATGACGGAAGGAAAGAACTGATGTTTAAAGTCCACTTTTAAACATATAT 4468
Db 1835 TAGGAACCATGACGGAAGGAAAGAACTGATGTTTAAAGTCCACTTTTAAACATATAT 1894
QY 4469 TTAATTTATGTAGATCTGTCAAGAGAACTTCCAAAGAGTTTATTAATTAACAGACT 4528
Db 1895 TTAATTTATGTAGATCTGTCAAGAGAACTTCCAAAGAGTTTATTAATTAACAGACT 1954
QY 4529 CTGTTGCAATGAATTA 4545
Db 1955 CTGTTGCAATGAATTA 1971

RESULT 15

AAK51624
ID AAK51624 standard; cDNA; 2120 BP.
XX
AC AAK51624;
XX
DT 06-NOV-2001 (first entry)
DE Human polynucleotide SEQ ID NO 169.
XX
KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukaemia;
KW nervous system disorder; arthritis; inflammation; ss.

XX Homo sapiens.
XX WO200157190-A2.
XX 09-AUG-2001.
XX 05-FEB-2001; 2001WO-US004098.
XX 03-FEB-2000; 2000US-00496914.
XX 27-APR-2000; 2000US-00560875.
XX 20-JUN-2000; 2000US-00598075.
XX 19-JUL-2000; 2000US-00620325.
XX 01-SEP-2000; 2000US-00654936.
XX 15-SEP-2000; 2000US-00663561.
XX 20-OCT-2000; 2000US-00693325.
XX 30-NOV-2000; 2000US-0078422.
XX (HYSE-) HYSEQ INC.
XX Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y;
XX Ma Y, Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;
XX Xue AJ, Yang Y, Wehrman T, Goodrich R;
XX WPI; 2001-476283/51.
XX P-PSDB; AAM78491.
XX Nucleic acids encoding polypeptides with cytokine-like activities, useful
XX in diagnosis and gene therapy.
XX Claim 1; Page 904-906; 6221pp; English.
XX The invention relates to polynucleotides (AAK51456-AAK53435) and the
XX encoded polypeptides (AAM78323-AAM80302) that exhibit activity relating to
XX cytokine, cell proliferation or cell differentiation or which may induce
XX production of other cytokines in other cell populations. The
XX polynucleotides and polypeptides are useful in gene therapy, vaccines or
XX peptide therapy. The polypeptides have various cytokine-like activities,
XX e.g. stem cell growth factor activity, haematopoiesis regulating
XX activity, tissue growth factor activity, immunomodulatory activity and
XX activin/inhibin activity and may be useful in the diagnosis and/or
XX treatment of cancer, leukaemia, nervous system disorders, arthritis and
XX inflammation. Note: Records for SEQ ID NO 2110 (AAK52581), 2111
XX (AAK52582) and 3666 (AAM80020) are omitted as the relevant pages from the
XX sequence listing were missing at the time of publication
XX SQ Sequence 2120 BP; 712 A; 390 C; 479 G; 539 T; 0 U; 0 Other;
Query Match 13.5%; Score 658.6; DB 4; Length 2120;
Best Local Similarity 99.3%; Pred. No. 4e-128;
Matches 672; Conservative 0; Mismatches 4; Indels 1; Gaps 1;
QY 3869 CAGGTTAAATCAGATCCAGAGAAACAGTGTCTTAAAGAGAGCGGTGGTGGTGA 3928
Db 1440 CTGGTTAAATCAGATCCAGAGAAACAGTGTCTTAAAGAGAGCGGTGGTGGTGA 1499
QY 3929 TAATAGATGTCTAGCAGCAATCCAAACGCGAGATCTACTGGGGTGGACAGTACACCTG 3988
Db 1500 TAATAGATGTCTAGCAGCAATCCAAACGCGAGATCTACTGGGGTGGACAGTACACCTG 1559
QY 3989 GGACATGGCAAGCATGGCACAGATGATGGTGTAGTATGGATGAATTTGGAAGGGTCTATG 4048
Db 1560 GGACATGGCAAGCATGGCACAGATGATGGTGTAGTATGGATGAATTTGGAAGGGTCTATG 1619
QY 4049 GTACTCAATGAGGAAGATGATGAAGATCAGGCCCTTCTTCCACAGCAATAGTCCCC 4108
Db 1620 GTACTCAATGAGGAAGATGATGAAGATCAGGCCCTTCTTCCACAGCAATAGTCCCC 1679
QY 4109 AATACGTAGATTTTGTCTCTGTATGTGACAAATTTTGTACATATGTTATTGGAA 4168
Db 1680 AATACGTAGATTTTGTCTCTGTATGTGACAAATTTTGTACATATGTTATTGGAA 1739
QY 4169 TTTTCTTTCATATPATATTCCTCTTAAACTCTCAAGCAGAGCGTGAAGTGTGACTTTTGG 4228

Db	1740	TTTTCTTTTCAATATATATTCCTTAAACTCTCAAGCAGACGTCAGTGTGACTTTTTG	1799
Qy	4229	AAAAAGTATAGGATATATTAATTAATTAATAGCAGATGATTTCTTTTGTCTTCTTCATT	4288
Db	1800	AAAAAGTATAGGATATATTAATTAATTAATAGCAGATGATTTCTTTTGTCTTCTTCATT	1859
Qy	4289	TCTCTTGCTCACCCCAAGTAACAAAAAGTATAGTTTTGACAGAGTTCGGTGTTCATAATT	4348
Db	1860	TCTCTTGCTCA-CCAGAGAGTAACAAAAGTATAGTTTTGACAGAGTTCGGTGTTCATAATT	1918
Qy	4349	TCAGTTCTAGTTGATTCGAGAAATTTCAAATAAGGAAGAGGGGTCTTTTATCCTTGTCTG	4408
Db	1919	TCAGTTCTAGTTGATTCGAGAAATTTCAAATAAGGAAGAGGGGTCTTTTATCCTTGTCTG	1978
Qy	4409	TAGGAAACCATGACGGAAGGAAGAACTGTGTTTAAAGTCCACTTTTAAACTATAT	4468
Db	1979	TAGGAAACCATGACGGAAGGAAGAACTGTGTTTAAAGTCCACTTTTAAACTATAT	2038
Qy	4469	TTATTTATGTAGGATCTGTCAAAGAAAACTTCCAAAAAGATTTATTAATTAACCCAGACT	4528
Db	2039	TTATTTATGTAGGATCTGTCAAAGAAAACTTCCAAAAAGATTTATTAATTAACCCAGACT	2098
Qy	4529	CTGTTGCAATAAGTTAA	4545
Db	2099	CTGTTGCAATAAAAAA	2115

Search completed: November 22, 2004, 17:23:20
Job time : 2173.38 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 22, 2004, 15:08:21 ; Search time 20190.7 Seconds
(without alignments)
11427.340 Million cell updates/sec

Title: US-10-017-724-5_COPY_4000_8878

Perfect score: 4879

Sequence: 1 tcagaatagtggtgatgag.....gattgcattgagccaagatc 4879

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4526729 seqs, 23644849745 residues

Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb.ba.*

2: gb.htg.*

3: gb.in.*

4: gb.om.*

5: gb.ov.*

6: gb.pat.*

7: gb.ph.*

8: gb.pl.*

9: gb.pr.*

10: gb.ro.*

11: gb.sts.*

12: gb.sy.*

13: gb.un.*

14: gb.vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4879	100.0	8878	6	I47706 Sequence 3
2	4879	100.0	8878	9	HUMFIBRB
3	4879	100.0	8878	11	G28617 SHGC-35791
4	4783	98.0	156442	9	AC107385 Homo sapi
5	4779.4	98.0	11956	9	AF388026
6	2332.4	47.8	330999	2	AC009583 Homo sapi
7	2092.4	42.9	330999	2	AC009583 Homo sapi
8	2081.2	42.7	4043	6	AX535001 Sequence
9	971.8	19.9	4967	6	CO414791
10	971.8	19.9	4967	6	AX285153 Sequence
11	696.2	14.3	223534	2	AC140369 Mus muscu
12	696.2	14.3	227600	10	AC138394
13	683	14.0	763	9	HUMFBRB
14	655.4	13.4	1792	9	BC036007 Homo sapi
15	625.4	13.4	1883	9	HUMFBRB
16	625.4	12.8	142856	2	AC143097
17	363.8	7.5	241740	2	AC111892 Rattus no
18	363.8	7.5	261110	2	AC105842 Rattus no
19	363.8	7.5	263108	2	AC112034 Rattus no

C	20	362.2	7.4	571	6	AX410537
C	21	338.2	6.9	375	6	AX410554
C	22	337	6.9	384	11	G25542
	23	336.4	6.9	37092	2	AC016278
	24	320.8	6.6	797	9	HUMFBRB
	25	288.2	5.9	1229	6	CO723670
	26	283.8	5.8	293	11	G05688
	27	283.8	5.8	293	11	G33709
C	28	264.4	5.4	357	6	AX408104
	29	246.2	5.0	251	9	AF200354S3
	30	246	5.0	261	10	RATFIBBA
	31	233	4.8	708	10	AF413205
	32	233	4.8	1602	10	BC031715
	33	233	4.8	1612	10	RNU05675
	34	230.2	4.7	1329	4	BTPIBR
	35	230	4.7	851	9	BC070245
	36	228.6	4.7	380	6	AX409357
	37	228	4.7	1007	10	RATFIBBB
	38	200.4	4.1	1763	5	CHKFIBBC
	39	181.4	3.7	1875	5	BC054326
	40	181.4	3.7	2390	5	XL019618
	41	180	3.7	230514	2	AC106233
	42	162	3.3	162	11	GS4120
C	43	157.8	3.2	181272	5	BX548256
C	44	157.2	3.2	190325	9	AL732327
C	45	156.6	3.2	3556	9	AF230338

ALIGNMENTS

RESULT 1	I47706	Sequence 3 from patent US 5639940.	8878 bp	DNA	linear	PAT 07-OCT-1997
LOCUS	I47706	Sequence 3 from patent US 5639940.	8878 bp	DNA	linear	PAT 07-OCT-1997
DEFINITION	I47706	Sequence 3 from patent US 5639940.	8878 bp	DNA	linear	PAT 07-OCT-1997
ACCESSION	I47706	Sequence 3 from patent US 5639940.	8878 bp	DNA	linear	PAT 07-OCT-1997
VERSION	I47706.1	GI:2471671	8878 bp	DNA	linear	PAT 07-OCT-1997
KEYWORDS	Unknown.	GI:2471671	8878 bp	DNA	linear	PAT 07-OCT-1997
SOURCE	Unknown.	GI:2471671	8878 bp	DNA	linear	PAT 07-OCT-1997
ORGANISM	Unknown.	GI:2471671	8878 bp	DNA	linear	PAT 07-OCT-1997
REFERENCE	1 (bases 1 to 8878)	GI:2471671	8878 bp	DNA	linear	PAT 07-OCT-1997
AUTHORS	Garner, I., Dallymple, M.L., Prunkard, D.E. and Foster, D.C.	GI:2471671	8878 bp	DNA	linear	PAT 07-OCT-1997
TITLE	Production of fibrinogen in transgenic animals	GI:2471671	8878 bp	DNA	linear	PAT 07-OCT-1997
JOURNAL	Patent: US 5639940-A 3 17-JUN-1997;	GI:2471671	8878 bp	DNA	linear	PAT 07-OCT-1997
FEATURES	Location/Qualifiers	GI:2471671	8878 bp	DNA	linear	PAT 07-OCT-1997
source	1. .8878	GI:2471671	8878 bp	DNA	linear	PAT 07-OCT-1997
ORIGIN	/organism="unknown"	GI:2471671	8878 bp	DNA	linear	PAT 07-OCT-1997
	/mol_type="unassigned DNA"	GI:2471671	8878 bp	DNA	linear	PAT 07-OCT-1997
Query Match	100.0%	Score 4879;	DB 6;	Length 8878;		
Best Local Similarity	100.0%	Pred. No. 0;				
Matches 4879;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;		
QY	1	TCAGAAATAGTCTGATGAGTTAAATAACAATGTGGAAGCTGTTTCCAGAGCTCTCTTT	60			
Db	4000	TCAGAAATAGTCTGATGAGTTAAATAACAATGTGGAAGCTGTTTCCAGAGCTCTCTTT	4059			
QY	61	CTTCTTTTCAGTACATGATTTTGTGGAAGACCTGTGCAAAAGAGGAGCAAGTAAT	120			
Db	4060	CTTCTTTTCAGTACATGATTTTGTGGAAGACCTGTGCAAAAGAGGAGCAAGTAAT	4119			
QY	121	AGGTAGATATCTTGTGCTTTCATTCGATTTTCAGCTATTAATTTGGAACCGTTAGAC	180			
Db	4120	AGGTAGATATCTTGTGCTTTCATTCGATTTTCAGCTATTAATTTGGAACCGTTAGAC	4179			
QY	181	TGCCACGAGATGATGTTGTGAGAGATTAACATTTCTGGTGTAGTGAATGAGATTCA	240			
Db	4180	TGCCACGAGATGATGTTGTGAGAGATTAACATTTCTGGTGTAGTGAATGAGATTCA	4239			
QY	241	TACGCTTTTGGGACCTTCCCTCCCTGCAACTTGCAGATAGCACTATTTCAGCTTTATTC	300			

2461 GCAGTTTCCAAAGATTTTACAGTTTACCTCTCATATTTATTTTCTTATTTGTTGTTCTATTT 2520
Db
6460 GCAGTTTCCAAAGATTTTACAGTTTACCTCTCATATTTATTTTCTTATTTGTTGTTCTATTT 6519
QY
2521 TAGAGCACCAATATATATCTAAATGGAATGGAACGGGATTCAGATATTTATTTTCAAAGT 2580
Db
6520 TAGAGCACCAATATATATCTAAATGGAATGGAACGGGATTCAGATATTTATTTTCAAAGT 6579
QY
2581 GACATTTATTTGTTGTTTAAATATATATCTCTTTTGTGTTCTCTCAACCAAGGATGGAC 2640
Db
6580 GACATTTATTTGTTGTTTAAATATATATCTCTTTTGTGTTCTCTCAACCAAGGATGGAC 6639
QY
2641 AGTGATTCAGAACCGTCAAGACGGTAGTGTGACITTTGGCAGGAATGGGATCCCATATAA 2700
Db
6640 AGTGATTCAGAACCGTCAAGACGGTAGTGTGACITTTGGCAGGAATGGGATCCCATATAA 6699
QY
2701 ACAGGGATTTGGAATGTTGCAACCAACACAGATGGGAAGAATTAATCTGFGGCTACCAGG 2760
Db
6700 ACAGGGATTTGGAATGTTGCAACCAACACAGATGGGAAGAATTAATCTGFGGCTACCAGG 6759
QY
2761 TAACGAACAGGCATCAAAATAAATCAATCTATTTGGAATGGGATTTTATTTAATTAA 2820
Db
6760 TAACGAACAGGCATCAAAATAAATCAATCTATTTGGAATGGGATTTTATTTAATTAA 6819
QY
2821 AAACATTCATTTGGAAGCCTGTTTGGCAGTTTAAAGAGGATTTCCCTGACAAAAATGT 2880
Db
6820 AAACATTCATTTGGAAGCCTGTTTGGCAGTTTAAAGAGGATTTCCCTGACAAAAATGT 6879
QY
2881 GGAAGCTAAAGTAAAGGAAGAAAGCAGTTTATTTAGTTTCCCAAAATTTATTTTGGTG 2940
Db
6880 GGAAGCTAAAGTAAAGGAAGAAAGCAGTTTATTTAGTTTCCCAAAATTTATTTTGGTG 6939
QY
2941 AGAGATTTTATTTGTTTCTTTTAGGTGAATATTTGGCTTGGAAATGATAAAATTTAGCC 3000
Db
6940 AGAGATTTTATTTGTTTCTTTTAGGTGAATATTTGGCTTGGAAATGATAAAATTTAGCC 6999
QY
3001 AGCTTACCAGATGGGACCCACAGAACTTTTGATAGAAATGAGGAGCTGGAAGGAGACA 3060
Db
7000 AGCTTACCAGATGGGACCCACAGAACTTTTGATAGAAATGAGGAGCTGGAAGGAGACA 7059
QY
3061 AAGTAAAGGCTCACTATGAGGATTCATGTTACAGAAATGAAGCCCAAAATACAGATCT 3120
Db
7060 AAGTAAAGGCTCACTATGAGGATTCATGTTACAGAAATGAAGCCCAAAATACAGATCT 7119
QY
3121 CAGTGAACAAATACAGAGAAACAGCGGTAATGCCCTCATGATGGAGCATCTCAGCTGA 3180
Db
7120 CAGTGAACAAATACAGAGAAACAGCGGTAATGCCCTCATGATGGAGCATCTCAGCTGA 7179
QY
3181 TGGGAGAAAACAGGACCATGACCATTCACACGGCATGTTCTTCAGCACGTATGACAGAG 3240
Db
7180 TGGGAGAAAACAGGACCATGACCATTCACACGGCATGTTCTTCAGCACGTATGACAGAG 7239
QY
3241 ACAATGACGGCTGGTATGTTGGGACTCTTTGCTCTGCTTTTAAAAATCACAATATTC 3300
Db
7240 ACAATGACGGCTGGTATGTTGGGACTCTTTGCTCTGCTTTTAAAAATCACAATATTC 7299
QY
3301 ATTACTCAGAATCATTAACAATATTTTAAATAGCTACCACTTCTCGGACATTTCTGTCA 3360
Db
7300 ATTACTCAGAATCATTAACAATATTTTAAATAGCTACCACTTCTCGGACATTTCTGTCA 7359
QY
3361 GCCACTGTCTTAAGCTCTTTATGCACTCACTCGAAAGCATTTTCAACTATAAGGTAGACATT 3420
Db
7360 GCCACTGTCTTAAGCTCTTTATGCACTCACTCGAAAGCATTTTCAACTATAAGGTAGACATT 7419
QY
3421 CTTATTTCTCATTTTACAGATGAGATTTAGAGAGATTAACGTGATTTGTCOAATGTCACACA 3480
Db
7420 CTTATTTCTCATTTTACAGATGAGATTTAGAGAGATTAACGTGATTTGTCOAATGTCACACA 7479
QY
3481 ACTACCCAGAGATAAACTAGAAATTTGACACAGTTTCTTCTGATATAGGACATTTAG 3540
Db
7480 ACTACCCAGAGATAAACTAGAAATTTGACACAGTTTCTTCTGATATAGGACATTTAG 7539
QY
3541 ATAAATACCTATATCTATATTTCTAAAGTGTGTGAAAACTTTTCATTTTCTTCCAG 3600

7540 ATAAATACCTATATCTATATTTCTAAAGTGTGTGAAAACTTTTCATTTTCTTCCAG 7599
QY
3601 GGTTCCTCGATATAAGGGTGTAAAAAGCTATTTTCCAGTATAAAGTAAACAACACAGT 3660
Db
7600 GGTTCCTCGATATAAGGGTGTAAAAAGCTATTTTCCAGTATAAAGTAAACAACACAGT 7659
QY
3661 CCCTAGATGGATTCGACAAAGGGCCAGTTATCTCTTTCTTCTATAGGACACAGGAG 3720
Db
7660 CCCTAGATGGATTCGACAAAGGGCCAGTTATCTCTCTTTCTTCTATAGGACACAGGAG 7719
QY
3721 GTCTTTGGTGTATTAAGTGTGACTATGTATAGCACCCAAAAGGAAAGACTACTGTGCACA 3780
Db
7720 GTCTTTGGTGTATTAAGTGTGACTATGTATAGCACCCAAAAGGAAAGACTACTGTGCACA 7779
QY
3781 CGAGTGTAGCAGTCTTTTATGGGTAAATCTGCAAAAAGTAACTTGGACCCAGTGTCTGT 3840
Db
7780 CGAGTGTAGCAGTCTTTTATGGGTAAATCTGCAAAAAGTAACTTGGACCCAGTGTCTGT 7839
QY
3841 TTCTTAATACGCCAAACACATTTTCTTCAGGTAAACATCAGATCCCAGAAAACAGTGT 3900
Db
7840 TTCTTAATACGCCAAACACATTTTCTTCAGGTAAACATCAGATCCCAGAAAACAGTGT 7899
QY
3901 CTAAAGAAAGACGGTGGTGTATTAAGTGTATAGATGTCTATGACCAATCCAAACGGCA 3960
Db
7900 CTAAAGAAAGACGGTGGTGTATTAAGTGTATAGATGTCTATGACCAATCCAAACGGCA 7959
QY
3961 GATACTACTGGGGTGGACAGTACACCTGGGACATGGCAAGCATGGCACAGATGATGTG 4020
Db
7960 GATACTACTGGGGTGGACAGTACACCTGGGACATGGCAAGCATGGCACAGATGATGTG 8019
QY
4021 TAGTATGATGAATTTGGAAGGGTCTATGTTCTCAATGAGGAGATGATGATGAAGATCA 4080
Db
8020 TAGTATGATGAATTTGGAAGGGTCTATGTTCTCAATGAGGAGATGATGATGAAGATCA 8079
QY
4081 GGCCCTCTTCCCAACAGCAATAGTCCCAATACGTAGATTTTGTCTCTCTGATGTGAC 4140
Db
8080 GGCCCTCTTCCCAACAGCAATAGTCCCAATACGTAGATTTTGTCTCTCTGATGTGAC 8139
QY
4141 AACATTTTGTACATATGTTATTTGGAAATTTTCTTTTCTATATATTTCTCTTAAACT 4200
Db
8140 AACATTTTGTACATATGTTATTTGGAAATTTTCTTTTCTATATATTTCTCTTAAACT 8199
QY
4201 CTCAGACAGCTGATGCTGCTTTTGAAGGATGATGATGATGATGATGATGATGATGATG 4260
Db
8200 CTCAGACAGCTGATGCTGCTTTTGAAGGATGATGATGATGATGATGATGATGATGATG 8259
QY
4261 CACATGATTTTCTTTTGTGTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 4320
Db
8260 CACATGATTTTCTTTTGTGTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 8319
QY
4321 AGTTTGTACAGATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 4380
Db
8320 AGTTTGTACAGATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 8379
QY
4381 AAGGAAGGGCTCTTTTATCTTCTGCTAGGAAACCATGACGGAAGGAAAGGAAAGTGTAT 4440
Db
8380 AAGGAAGGGCTCTTTTATCTTCTGCTAGGAAACCATGACGGAAGGAAAGGAAAGTGTAT 8439
QY
4441 GTTTAAAAAGTCCACTTTTAAAAACTATATTTATTTATGATGATCTGTCAAGAAAACTTC 4500
Db
8440 GTTTAAAAAGTCCACTTTTAAAAACTATATTTATTTATGATGATCTGTCAAGAAAACTTC 8499
QY
4501 CAAAAGATTTATTAATTAACCCAGACTCTGTTGCAATAGTAAATGTTTCTGTTTGTG 4560
Db
8500 CAAAAGATTTATTAATTAACCCAGACTCTGTTGCAATAGTAAATGTTTCTGTTTGTG 8559
QY
4561 TAAATCCACATTTCAATGATGTTGCTTTGCACTTTGTAAGGAGGAGGAGCGTTTCAAC 4620
Db
8560 TAAATCCACATTTCAATGATGTTGCTTTGCACTTTGTAAGGAGGAGGAGCGTTTCAAC 8619
QY
4621 CTCAATAGCTAAATAAACCGGTCTTTGAATATTTGAAGATTTTAAAAATCTGACTTAGGACG 4680

781 TAAACAACTATTTTCAATTTACTGACACAACTCTTTAGCGTATACCTATGTTAAATTA 840
Db
4780 TAAACAACTATTTTCAATTTACTGACACAACTCTTTAGCGTATACCTATGTTAAATTA 4839
QY
841 CTAGTATGTTGGTTAGGATTTATGTTAAATTTGATATGTCATCGCGCAAAATCAATTTCCAC 900
Db
4840 CTAGTATGTTGGTTAGGATTTATGTTAAATTTGATATGTCATCGCGCAAAATCAATTTCCAC 4899
QY
901 TAAATATGACTAT 960
Db
4900 TAAATATGACTAT 4959
QY
961 AGAAATCAAAATTTGATATAGTTTAAATACATTTAGTTTATGAGGCAAAATGCTTAACTATT 1020
Db
4960 AGAAATCAAAATTTGATATAGTTTAAATACATTTAGTTTATGAGGCAAAATGCTTAACTATT 5019
QY
1021 TCTACATAATTTCAATTTTCCAGATTAATGAAATGATGCAATGATGTAATTTCTCAGAACT 1080
Db
5020 TCTACATAATTTCAATTTTCCAGATTAATGAAATGATGCAATGATGTAATTTCTCAGAACT 5079
QY
1081 GGAAAGACCAAT 1140
Db
5080 GGAAAGACCAAT 5139
QY
1141 TGTGCTTCGTTCAATTCCTGGAACCTGAGAACCAAAATACAAAGTTAGAAATCTCATGT 1200
Db
5140 TGTGCTTCGTTCAATTCCTGGAACCTGAGAACCAAAATACAAAGTTAGAAATCTCATGT 5199
QY
1201 CTCAGCTCAAAATGGAATATTTGCGACCCCATGCACTGTCAGTTGCAATTTCTGTTGT 1260
Db
5200 CTCAGCTCAAAATGGAATATTTGCGACCCCATGCACTGTCAGTTGCAATTTCTGTTGT 5259
QY
1261 GTCTGCAAGGTAACTGATTCATTAACATATTTTATGAGAGTTCCAGAAAGTCTCACAC 1320
Db
5260 GTCTGCAAGGTAACTGATTCATTAACATATTTTATGAGAGTTCCAGAAAGTCTCACAC 5319
QY
1321 ACCAAAAATAAGAGAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAAC 1380
Db
5320 ACCAAAAATAAGAGAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAAC 5379
QY
1381 ATAATGACATTAAGTAT 1440
Db
5380 ATAATGACATTAAGTAT 5439
QY
1441 GGTGCAATTTGCTGGTTTATGAGCAGGATGGGGCACAGGTATCCAGGGTGGCTGAT 1500
Db
5440 GGTGCAATTTGCTGGTTTATGAGCAGGATGGGGCACAGGTATCCAGGGTGGCTGAT 5499
QY
1501 GTGTGCACTGCGTACTGCTTGAACAGATGGCAGAACCAACAGATAGATGTTAGAGTTTC 1560
Db
5500 GTGTGCACTGCGTACTGCTTGAACAGATGGCAGAACCAACAGATAGATGTTAGAGTTTC 5559
QY
1561 TCCATTTTGTGTTCTGGAGCTCATGATATTCAGGACACAAAGGTGGAGAGAGC 1620
Db
5560 TCCATTTTGTGTTCTGGAGCTCATGATATTCAGGACACAAAGGTGGAGAGAGC 5619
QY
1621 TTTGTTTCATCTTTAGCAGATAAAGCTCTCAAACTGGTTGACCTTACTTAAAGTAAA 1680
Db
5620 TTTGTTTCATCTTTAGCAGATAAAGCTCTCAAACTGGTTGACCTTACTTAAAGTAAA 5679
QY
1681 ATGAAATCTAATATTTGTTATATATTTTCAAGGTCTATATATACACACTCTTTAGTA 1740
Db
5680 ATGAAATCTAATATTTGTTATATATTTTCAAGGTCTATATATACACACTCTTTAGTA 5739
QY
1741 ACTTATGTAATGTTATTTTAAAGAAATTTGTTAGTAAATACAAAGTAAATTTATGTCATAAAC 1800
Db
5740 ACTTATGTAATGTTATTTTAAAGAAATTTGTTAGTAAATACAAAGTAAATTTATGTCATAAAC 5799
QY
1801 CCTTGAACATAATGTTGTTCTTACATTTTGCAGAAATGTTAGGAAATTTATTTTGGTG 1860
Db
5800 CCTTGAACATAATGTTGTTCTTACATTTTGCAGAAATGTTAGGAAATTTATTTTGGTG 5899

QY
1861 GAAACATCTGAATGTATCTCATTTCAACCTGACAGTTCTGTCAAAACCGTATAGATATAC 1920
Db
5860 GAAACATCTGAATGTATCTCATTTCAACCTGACAGTTCTGTCAAAACCGTATAGATATAC 5919
QY
1921 TGTGACATGAATACAGAAATGAGGTAAAGCTTTTCGACAGTTTGTTCACCTGTTGATCTGT 1980
Db
5920 TGTGACATGAATACAGAAATGAGGTAAAGCTTTTCGACAGTTTGTTCACCTGTTGATCTGT 5979
QY
1981 AATTATTTGATACCGTAAATTCGACGAAACAAAGCCAGGTGTGTGGCTCATACCTGT 2040
Db
5980 AATTATTTGATACCGTAAATTCGACGAAACAAAGCCAGGTGTGTGGCTCATACCTGT 6039
QY
2041 AATTTCCAGCACCTTGGGAGGCCAAAGTGGCTGTATAGCTTTGAGCTTAGGAGTTGAACT 2100
Db
6040 AATTTCCAGCACCTTGGGAGGCCAAAGTGGCTGTATAGCTTTGAGCTTAGGAGTTGAACT 6099
QY
2101 AGCTGGGCAACATTAATGAGACCCCTAACTCTCAAAAAAATAAATAAATAAATAAATAA 2160
Db
6100 AGCTGGGCAACATTAATGAGACCCCTAACTCTCAAAAAAATAAATAAATAAATAAATAA 6159
QY
2161 AAAAAATCAGCTGTGTGTTGGTAGTATGTCCTGTAGTCCAGCTATCCAGGAGGCTGAGA 2220
Db
6160 AAAAAATCAGCTGTGTGTTGGTAGTATGTCCTGTAGTCCAGCTATCCAGGAGGCTGAGA 6219
QY
2221 TGGGAGATCACCTGAGCCCAACCACTGGAGTCTTGATCATGCTACTGAACTGTAGCCTGG 2280
Db
6220 TGGGAGATCACCTGAGCCCAACCACTGGAGTCTTGATCATGCTACTGAACTGTAGCCTGG 6279
QY
2281 GCAACAGAGGATGAGTGTGATCTCTCAAAAAAATAAATAAATAAATAAATAAATAAATAA 2340
Db
6280 GCAACAGAGGATGAGTGTGATCTCTCAAAAAAATAAATAAATAAATAAATAAATAAATAA 6339
QY
2341 AAGACTTAGCTCTAATCATCTAATGAGTGTGATGAGGAGTAAATTTGATGTTGAAATTCAC 2400
Db
6340 AAGACTTAGCTCTAATCATCTAATGAGTGTGATGAGGAGTAAATTTGATGTTGAAATTCAC 6399
QY
2401 CTGATATTTTAAAGTTTATAAATAATCTATAATTTCAAAATTTGGGGTAAAGTAAAGCATT 2460
Db
6400 CTGATATTTTAAAGTTTATAAATAATCTATAATTTCAAAATTTGGGGTAAAGTAAAGCATT 6459
QY
2461 GCAGTTTCCAAAGATTTTCAAGTTTACCTCTCATATTTTATTTTCTTATTTGTTCTATTT 2520
Db
6460 GCAGTTTCCAAAGATTTTCAAGTTTACCTCTCATATTTTATTTTCTTATTTGTTCTATTT 6519
QY
2521 TAGAGCACCAATATATATATGAAATGGAATGGAATGGAATGGAATGGAATGGAATGGAAT 2580
Db
6520 TAGAGCACCAATATATATGAAATGGAATGGAATGGAATGGAATGGAATGGAATGGAAT 6579
QY
2581 GACATTTATTTGCTGTTGTTAATATATATGCTCTTTTCTGTCGACCAACCAAGGATGGAC 2640
Db
6580 GACATTTATTTGCTGTTGTTAATATATGCTCTTTTCTGTCGACCAACCAAGGATGGAC 6639
QY
2641 AGTGATTCAGAACCGTCAAGCGGTAGTGTGACTTTTGGCAGAAATGGGATCCATATAA 2700
Db
6640 AGTGATTCAGAACCGTCAAGCGGTAGTGTGACTTTTGGCAGAAATGGGATCCATATAA 6699
QY
2701 ACAGGATTTGGAATTTGCAACCAACAGATGGAAGAAATTTACTGTGGCTTACCAGG 2760
Db
6700 ACAGGATTTGGAATTTGCAACCAACAGATGGAAGAAATTTACTGTGGCTTACCAGG 6759
QY
2761 TAAACGAAAGGATCAAAATAAATCATTTTATTTGAAATGGATTTTATTTTAAATTA 2820
Db
6760 TAAACGAAAGGATCAAAATAAATCATTTTATTTGAAATGGATTTTATTTTAAATTA 6819
QY
2821 AAAACATTTCTTTGGAAGCTTTTATAGGAGTAAAGAGGAGTTTCTTGCAAAAAATCT 2880
Db
6820 AAAACATTTCTTTGGAAGCTTTTATAGGAGTAAAGAGGAGTTTCTTGCAAAAAATCT 6879
QY
2881 GGAAGCTAAAGTAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2940
Db
6880 GGAAGCTAAAGTAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 6939
QY
2941 AGAGATTTTATTTTCTTTTCTTTTAGGTGAATATGCTTGGAAATGATAAATAATAGCC 3000

JOURNAL
COMMENT

Unpublished (2000)

Contact: Michael Olivier, David R. Cox
Stanford Human Genome Center
Stanford University School of Medicine
4005 Miranda Ave. 2nd Fl., Palo Alto, CA 94025, USA
Tel: (650) 320-5800
Fax: (650) 320-5801
Email: olivier@shgc.stanford.edu
Primer A: TCCCAATACAGTATTTTTCG
Primer B: ACTTTGTACTTCTTGGGTGAC
STS size: 215
PCR Profile:

Initial incubation: 94 degrees C for 90 seconds

Denaturation: 94 degrees C for 15 seconds
Annealing: 62 degrees C for 23 seconds
Polymerization: 72 degrees C for 30 seconds
PCR Cycles: 30
Thermal Cycler: Perkin Elmer 9600

Protocol:

Template: 25 ng
Primer: each 1 uM
dNTPs: each 200 uM
Tag Polymerase: 0.05 units/ul
Total Vol: 10 ul

Buffer: MgCl2: 2.5 mM
KCl: 50 mM
Tris-HCl: 20 mM
pH: 8.3

Prepared with primer pairs provided by Sandoz, derived from M64983
-- Washington University/Merck EST sequence.

FEATURES
source

Location/Qualifiers
1..8878
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/map="4"
/clone_lib="Human"
8103..8317
8103..8124
STS
primer_bind
primer_bind
complement(8294..8317)
ORIGIN

Query Match 100.0%; Score 4879; DB 11; Length 8878;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 4879; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TCAGAAATAGTGTGATGAGTTAAATAACAATGTGGAAGCTGTTCCCGAGCTCTCTTT 60
DB 4000 TCAGAAATAGTGTGATGAGTTAAATAACAATGTGGAAGCTGTTCCCGAGCTCTCTTT 4059
QY 61 CTTCTCTTCAGTACATGATTTTCTGCTGAAAGACCTGTGGCAAGAGGAGGAGCAAGTAA 120
DB 4060 CTTCTCTTCAGTACATGATTTTCTGCTGAAAGACCTGTGGCAAGAGGAGGAGCAAGTAA 4119
QY 121 AAGGTAGATATCTTGTGCTTCCATTCGATTTTCAGCTTATATAAATTTGGAACCGTTAGAC 180
DB 4120 AAGGTAGATATCTTGTGCTTCCATTCGATTTTCAGCTTATATAAATTTGGAACCGTTAGAC 4179
QY 181 TCCGAGGAATGCATGTTGTGAGAGATTAACATTTCTGGGTAGTGAATGATTTCA 240
DB 4180 TCCGAGGAATGCATGTTGTGAGAGATTAACATTTCTGGGTAGTGAATGATTTCA 4239
QY 241 TACGCTTTTGGGACCTTCCCTGCAACTTTCAGATAAGCACTATTTCAGCTCTTATTC 300
DB 4240 TACGCTTTTGGGACCTTCCCTGCAACTTTCAGATAAGCACTATTTCAGCTCTTATTC 4299
QY 301 CAGTCTGACATCAGCAAGTGTGATTTTCTATGAAAATTTCTACTATGATCTCTTATTTTA 360
DB 5380 ATAATGACATTACATACATCATATAAATATCTTTAGCCAGTTGTGTTTGGACTGGCCT 5439

DB 4300 CAGTCTGACATCAGCAAGTGTGATTTTCTATGAAAATTTCTACTATGACTCCTTATTTTA 4359
QY 361 AGTATACAGAAAACCTTGTGACTCAGAGATAATATTTTACAGAGTGGAAAAACCCCTAG 420
DB 4360 AGTATACAGAAAACCTTGTGACTCAGAGATAATATTTTACAGAGTGGAAAAACCCCTAG 4419
QY 421 CATTTATAGTTTTTAACATTTGAGGTTTTTGAATGAGAGAGTTATCCATAATATATTTCAATT 480
DB 4420 CATTTATAGTTTTTAACATTTGAGGTTTTTGAATGAGAGAGTTATCCATAATATATTTCAATT 4479
QY 481 GTGTTGTGATAATGACACCTAACCTGTGAATCTTTGAGGTGAGAAATGTTGAGTGTGTTG 540
DB 4480 GTGTTGTGATAATGACACCTAACCTGTGAATCTTTGAGGTGAGAAATGTTGAGTGTGTTG 4539
QY 541 ACTTGTGTGTCAGGAAAACAGCTAGTGTGCTGAGCTGGGACAGGCATCTCAGTGAGTAGCA 600
DB 4540 ACTTGTGTGTCAGGAAAACAGCTAGTGTGCTGAGCTGGGACAGGCATCTCAGTGAGTAGCA 4599
QY 601 TACCCACAGTTGGAATTTTTCAGAGAAATCAAGGAATCAAGGAATCTTATAAATTTTCA 660
DB 4600 TACCCACAGTTGGAATTTTTCAGAGAAATCAAGGAATCAAGGAATCTTATAAATTTTCA 4659
QY 661 AGGTTCTGCTATATCTTATGTGAAATGGAATAAATCAAGCATATCCACTCTGTAAGAT 720
DB 4660 AGGTTCTGCTATATCTTATGTGAAATGGAATAAATCAAGCATATCCACTCTGTAAGAT 4719
QY 721 TGAACCTTCTCAGATGGAAGACCCCAATCTGCTTTCTCTCTTTCCCTCACCAGAAAA 780
DB 4720 TGAACCTTCTCAGATGGAAGACCCCAATCTGCTTTCTCTCTTTCCCTCACCAGAAAA 4779
QY 781 TAAACAAACCTTATTTTACTTGGACACACATCTTTAGCGGTATACCTATGTAATTTA 840
DB 4780 TAAACAAACCTTATTTTACTTGGACACACATCTTTAGCGGTATACCTATGTAATTTA 4839
QY 841 CTAGTATGCTGTGTTAGGATTTTATGTTAAATTTGATATGTCATGCGCAAAATCATTTCCAC 900
DB 4840 CTAGTATGCTGTGTTAGGATTTTATGTTAAATTTGATATGTCATGCGCAAAATCATTTCCAC 4899
QY 901 TAAATATGACTATATATCATTAATCTGCTTGGTGAATGAGTCTGAGTGTGTTAATGTTATCTC 960
DB 4900 TAAATATGACTATATATCATTAATCTGCTTGGTGAATGAGTCTGAGTGTGTTAATGTTATCTC 4959
QY 961 AGAAAATCAAAATGTTATAGTTAAATACATTAATGTTTATGAGCAAAATGCTAACTATT 1020
DB 4960 AGAAAATCAAAATGTTATAGTTAAATACATTAATGTTTATGAGCAAAATGCTAACTATT 5019
QY 1021 TCTACATAATTTTCAATTTTCCAGATAATGAAAATGATGTAATGAGTACTCTCTCAGAACT 1080
DB 5020 TCTACATAATTTTCAATTTTCCAGATAATGAAAATGATGTAATGAGTACTCTCTCAGAACT 5079
QY 1081 GGAAAAGCAACCAATTTATATATAGATGAGACTGTGTAATGCAATATCCCAACTAACCTTCG 1140
DB 5080 GGAAAAGCAACCAATTTATATATAGATGAGACTGTGTAATGCAATATCCCAACTAACCTTCG 5139
QY 1141 TGTGCTTGTCTTCAATCTTGGAAAACTGAGAGCAAAATACAAAAGTTAGAAATCTGATCT 1200
DB 5140 TGTGCTTGTCTTCAATCTTGGAAAACTGAGAGCAAAATACAAAAGTTAGAAATCTGATCT 5199
QY 1201 CTCAGCTCAAAATGGAATATTTGTCGACCCCTGCACTGTGCTGCAATATTTCTGTGCT 1260
DB 5200 CTCAGCTCAAAATGGAATATTTGTCGACCCCTGCACTGTGCTGCAATATTTCTGTGCT 5259
QY 1261 GTCTGGCAAGGTAACTGATTCATTAACATATTTTATAGAGGTTCCAGAGAACTCACAC 1320
DB 5260 GTCTGGCAAGGTAACTGATTCATTAACATATTTTATAGAGGTTCCAGAGAACTCACAC 5319
QY 1321 ACCAAAAATAGAGAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAAC 1380
DB 5320 ACCAAAAATAGAGAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAAC 5379
QY 1381 ATAATGACATTACATACATCATATAAATATCTTTAGCCAGTTGTGTTTGGACTGGCCT 1440
DB 5380 ATAATGACATTACATACATCATATAAATATCTTTAGCCAGTTGTGTTTGGACTGGCCT 5439

This sequence was finished as follows unless otherwise noted:
all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:

Mapping information for this clone was provided by Dr. John D. McPherson, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu/gsc>

SOURCE INFORMATION:

The RPCI-11 human BAC library was made from the blood of one male donor, as described by Osoegawa, K., Woon, P.Y., Zhao, B., Frengen, E., Tatenio, M., Catanese, J.J. and de Jong, P.J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. Genomics 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong and coworkers at <http://www.chori.org>
VECTOR: pBACe3.6

NEIGHBORING SEQUENCE INFORMATION:

The clone sequenced to the left is RP13-58007; the clone sequenced to the right is RP11-216A23, 2000 bp overlap. Actual start of this clone is at base position 1 of RP11-158C21; actual end is at base position 32435 of RP11-216A23.

Polymorphisms have been identified between AC107385 and AC092616.

FEATURES

source

```
1. 156442
   Location/Qualifiers
     /organism="Homo sapiens"
     /mol_type="genomic DNA"
     /db_xref="taxon:9606"
     /chromosome="4"
     /map="4"
     /clone="RP11-158C21"
     /clone_lib="RPCI-11"
```

```
misc_feature
1. .646
   /note="similar to Homo sapiens EST BF205005
   (NID:gl11098591)"

misc_feature
1. .615
   /note="similar to Homo sapiens EST BF307717
   (NID:gl1254872)"

misc_feature
1. .535
   /note="similar to Homo sapiens EST BG394804
   (NID:gl13288252)"

misc_feature
1. .480
   /note="similar to Homo sapiens EST BI868734
   (NID:gl5042407)"

misc_feature
1. .417
   /note="similar to Homo sapiens EST BE312794
   (NID:gl131943)"

misc_feature
1. .337
   /note="similar to Homo sapiens EST BF237442
   (NID:gl1151360)"

misc_feature
1. .89
   /note="similar to Mus musculus EST BG093388
   (NID:gl12575951) ut70h03.y1"

misc_feature
1. .75
   /note="similar to Homo sapiens EST AI637897 (NID:ga690131)
   tt11f10.x1"

misc_feature
32. .765
   /note="similar to Homo sapiens EST BG755962
   (NID:gl14066615)"

misc_feature
66. .666
   /note="similar to Homo sapiens EST AL522055
   (NID:gl12785548)"

misc_feature
190. .565
   /note="similar to Mus musculus EST BG093388
```

```
misc_feature
(NID:gl12575951) ut70h03.y1"
220. .765
   /note="similar to Homo sapiens EST BM015896
   (NID:gl16530250)"

misc_feature
223. .798
   /note="similar to Mus musculus EST BI854901
   (NID:gl15995648)"

misc_feature
253
   /note="similar to Homo sapiens EST BG251146
   (NID:gl12760962)"

misc_feature
289. .765
   /note="similar to Homo sapiens EST AI912697 (NID:g5632552)
   we12f12.x1"

misc_feature
305. .765
   /note="similar to Homo sapiens EST W79609 (NID:gl1390880)
   zd81b07.r1"

misc_feature
335. .560
   /note="similar to Homo sapiens EST AW938856
   (NID:gl114302)"

misc_feature
342. .765
   /note="similar to Homo sapiens EST AI637897 (NID:ga690131)
   tt11f10.x1"

misc_feature
364. .570
   /note="similar to Homo sapiens EST AA428438 (NID:g2110249)
   zv50a04.r1"

misc_feature
497. .836
   /note="similar to EST BM676074 (NID:gl18985970)"

misc_feature
499. .765
   /note="similar to Homo sapiens EST BF037254
   (NID:gl0745479)"

misc_feature
507. .765
   /note="similar to Homo sapiens EST N56492 (NID:gl199340)"

misc_feature
540. .798
   /note="similar to Homo sapiens EST AL565079
   (NID:gl12916097)"

misc_feature
573. .798
   /note="similar to Homo sapiens EST BE395713
   (NID:gl9341078)"

misc_feature
607. .798
   /note="similar to Homo sapiens EST BG149868
   (NID:gl12661898)"

misc_feature
620. .798
   /note="similar to Homo sapiens EST BG251146
   (NID:gl12760962)"

misc_feature
695. .798
   /note="similar to Homo sapiens EST AU150237
   (NID:gl1011758)"

misc_feature
738. .798
   /note="similar to Homo sapiens EST AA284866 (NID:gl1927608)
   zt23a04.s1"

misc_feature
786. .794
   /note="match to EST AI669845 (NID:g4834619) wb85e08.x1"

misc_feature
811. .1179
   /note="similar to Homo sapiens EST W46662 (NID:gl131290)
   zc33d04.r1"

misc_feature
816. .1200
   /note="similar to Homo sapiens EST AL565079
   (NID:gl12916097)"

misc_feature
816. .1179
   /note="similar to Homo sapiens EST BG251146
   (NID:gl12760962)"

misc_feature
816. .1176
   /note="similar to Homo sapiens EST BE395713
   (NID:gl9341078)"

misc_feature
816. .1148
   /note="similar to Homo sapiens EST AA284866 (NID:gl1927608)
   zt23a04.s1"

misc_feature
816. .1148
   /note="similar to Homo sapiens EST AU150237
   (NID:gl1011758)"

misc_feature
816. .1103
   /note="similar to Homo sapiens EST AU150237
   (NID:gl1011758)"
```

Query Match

98.0%; Score 4783; DB 9; Length 156442;

QY 2161 AAAAAATCAGCTGTGTGGTAGTATGCTGTAGTCCAGCTATCCAGAGGCTGAGA 2220
Db |||||
QY 44447 AAAAAATCAGCTGTGTGGTAGTATGCTGTAGTCCAGCTATCCAGAGGCTGAGA 44506
Db |||||
QY 2221 TGGAGATCAGCTGTAGCCCAACAACTGGAGTCTTGTATCATCTACTGAATGTAGCTGG 2280
Db |||||
QY 44507 TGGAGATCAGCTGTAGCCCAACAACTGGAGTCTTGTATCATCTACTGAATGTAGCTGG 44566
Db |||||
QY 2281 GCAACAGAGGATGTAGATCCTGTCTCAAAAAAATAATTAATTAAGAGCCAGGAAC 2340
Db |||||
QY 44567 GCAACAGAGGATGTAGATCCTGTCTCAAAAAAATAATTAATTAAGAGCCAGGAAC 44626
Db |||||
QY 2341 AAGACTTGTCTTAACATCTAACATGTGACAAAGAGTAAATTTGATGTGGAATTTCAAC 2400
Db |||||
QY 44627 AAGACTTGTCTTAACATCTAACATGTGACAAAGAGTAAATTTGATGTGGAATTTCAAC 44685
Db |||||
QY 2401 CTGATATTTT-AAAAGTTATAAATAATCTATTAATTCACAATTTGGGGTAAAGATCACT 2459
Db |||||
QY 44686 CTGATATTTTAAAAGTTATAAATAATCTATTAATTCACAATTTGGGGTAAAGATCACT 44745
Db |||||
QY 2460 TGCAGTTTCCAAAGATTTTACAAGTTTACCTCTCATATTTATTTCTTATTTGTGCTATT 2519
Db |||||
QY 44746 TGCAGTTTCCAAAGATTTTACAAGTTTACCTCTCATATTTATTTCTTATTTGTGCTATT 44805
Db |||||
QY 2520 TTAGAGCACCAATATATATACTAAATGGAATGGACAGGGGATTCAGATATTTTCAAAG 2579
Db |||||
QY 44806 TTAGAGCACCAATATATATACTAAATGGAATGGACAGGGGATTCAGATATTTTCAAAG 44865
Db |||||
QY 2580 TGACATATTTGCTGTGGTTAATATATATGCTCTTTTGTGTCGACCAACAAAGATGGA 2639
Db |||||
QY 44866 TGACATATTTGCTGTGGTTAATATATATGCTCTTTTGTGTCGACCAACAAAGATGGA 44925
Db |||||
QY 2640 CAGTCAATTCAGAACCGTCAAGACGCTAGTGTGACTTTGGCAGGAATGGGATCCATATA 2699
Db |||||
QY 44926 CAGTCAATTCAGAACCGTCAAGACGCTAGTGTGACTTTGGCAGGAATGGGATCCATATA 44985
Db |||||
QY 2700 AACAGGATTTGGAATGTTGCAACCAACACAGATGGGAAGAAATTTACTGTGGCTTACCAG 2759
Db |||||
QY 44986 AACAGGATTTGGAATGTTGCAACCAACACAGATGGGAAGAAATTTACTGTGGCTTACCAG 45045
Db |||||
QY 2760 GTAAAGAACAGGATGCAAAATAAATAATCAATCTTATTTGAAATGGGATTTTTTTAATTA 2819
Db |||||
QY 45046 GTAAAGAACAGGATGCAAAATAAATAATCAATCTTATTTGAAATGGGATTTTTTTAATTA 45105
Db |||||
QY 2820 AAAACATTTGTTGGAGCTGTTTGGAGCTGTTTGGAGAGTTTAAAGAGTCTCTGCAAAAATG 2879
Db |||||
QY 45106 AAAACATTTGTTGGAGCTGTTTGGAGCTGTTTGGAGAGTTTCTGCAAAAATG 45165
Db |||||
QY 2880 TGGAGCTTAAAGATAAGGGAAGAAAGGAGGAGTTTGTAGTTTCCCAAAATTTTATTTTGGT 2939
Db |||||
QY 45166 TGGAGCTTAAAGATAAGGGAAGAAAGGAGGAGTTTGTAGTTTCCCAAAATTTTATTTTGGT 45225
Db |||||
QY 2940 GAGAGATTTTATTTTGTCTTTTGTAGTGAATATGCTGTTGGAATGATTAATTAATAGC 2999
Db |||||
QY 45226 --GAGATTTTATTTTGTCTTTTGTAGTGAATATGCTGTTGGAATGATTAATTAATAGC 45283
Db |||||
QY 3000 CAGCTTACCAGGATGGAGCCCAAGAACTTTTGTAGTAAGTGGAGCTGGAAAGGAGAC 3059
Db |||||
QY 45284 CAGCTTACCAGGATGGAGCCCAAGAACTTTTGTAGTAAGTGGAGCTGGAAAGGAGAC 45343
Db |||||
QY 3060 AAAGTAAAGGCTCAGTATGGAGGATTCAGTGTACAGAAATGAAGCCCAACAAATACCAGATC 3119
Db |||||
QY 45344 AAAGTAAAGGCTCAGTATGGAGGATTCAGTGTACAGAAATGAAGCCCAACAAATACCAGATC 45403
Db |||||
QY 3120 TCAGTGAACAAATACAGAGAACAGCCGTTATGCTTATGCTTATGATGGAGCATCTCAGCTG 3179
Db |||||
QY 45404 TCAGTGAACAAATACAGAGAACAGCCGTTATGCTTATGCTTATGATGGAGCATCTCAGCTG 45463
Db |||||
QY 3180 ATGGGAGAAAAAGGACCATGACATTCACAAAGGATGTTCTTCAGACGATGATGACAGA 3239
Db |||||
QY 45464 ATGGGAGAAAAAGGACCATGACATTCACAAAGGATGTTCTTCAGACGATGATGACAGA 45523
Db |||||

QY 3240 GACAATGACGCTGGTATGTGTGGACATCTTTGTCTCTGTCTTTAAAAATCACTAATAT 3299
Db |||||
QY 45524 GACAATGACGCTGGTATGTGTGGACATCTTTGTCTCTGTCTTTAAAAATCACTAATAT 45583
Db |||||
QY 3300 CATTACTCAGNAATCAATAAAGATATTTTAAATAGTACCACTCTCTGGGACTTACTGTG 3359
Db |||||
QY 45584 CATTACTCAGNAATCAATAAAGATATTTTAAATAGTACCACTCTCTGGGACTTACTGTG 45643
Db |||||
QY 3360 AGCCACTGTCTTAAGCTCTTTTATGATCATCTCGAAAGCATTTCAACTATAAGGTAGACAT 3419
Db |||||
QY 45644 AGCCACTGTCTTAAGCTCTTTTATGATCATCTCGAAAGCATTTCAACTATAAGGTAGACAT 45703
Db |||||
QY 3420 TCTTATTTCTCAATTTTACAGATGAGATTTAGAGAGATTTACGTGATTTGTCOAATGTCAAC 3479
Db |||||
QY 45704 TCTTATTTCTCAATTTTACAGATGAGATTTAGAGAGATTTACGTGATTTGTCOAATGTCAAC 45763
Db |||||
QY 3480 AACTACCCAGAGATAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAA 3539
Db |||||
QY 45764 AACTACCCAGAGATAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAA 45823
Db |||||
QY 3540 GATAAATACCTATATCTCTATATTTCTAAAGTGTGTGGAACCTTTTCAATTTTCCCA 3599
Db |||||
QY 45824 GATAAATACCTATATCTCTATATTTCTAAAGTGTGTGGAACCTTTTCAATTTTCCCA 45883
Db |||||
QY 3600 GGGTTCTCTGATACATAAGGGTTGTAAGAGCTTATTTCCAGTATTAAGTAAACACAG 3659
Db |||||
QY 45884 GGGTTCTCTGATACATAAGGGTTGTAAGAGCTTATTTCCAGTATTAAGTAAACACAG 45943
Db |||||
QY 3660 TCCCTAGATGAGATT-GCCACAAAGCCAGTTATCTCTCTTTCTTGTCTATAGGGCAGAG 3718
Db |||||
QY 45944 TCCCTAGATGAGATTGCGCACAAAGCCAGTTATCTCTCTTTCTTGTCTATAGGGCAGAG 46003
Db |||||
QY 3719 AGGTCTTTGGTGTATTTAGTGTGACTCTATGTATAGCACCCAAAGAAAGACTTACTGTGCA 3778
Db |||||
QY 46004 AGGTCTTTGGTGTATTTAGTGTGACTCTATGTATAGCACCCAAAGAAAGACTTACTGTGCA 46063
Db |||||
QY 3779 CACGAGTGTAGAGTCTTTTATGGGTATCTGCAAAAGCTAACTTGACCCCGTATGTTCT 3838
Db |||||
QY 46064 CACGAGTGTAGAGTCTTTTATGGGTATCTGCAAAAGCTAACTTGACCCCGTATGTTCT 46123
Db |||||
QY 3839 GTTTCTAATAACGCCAAACACATTTTCTTTCAGGTTTAACTCAGATCCAGAAAACAGTG 3898
Db |||||
QY 46124 GTTTCTAATAACGCCAAACACATTTTCTTTCAGGTTTAACTCAGATCCAGAAAACAGTG 46183
Db |||||
QY 3899 TTTCTAAAGAACGCGTGTGGTGTATTAATAGATGTCTATGATGATGATGATGATGATGATG 3958
Db |||||
QY 46184 TTTCTAAAGAACGCGTGTGGTGTATTAATAGATGTCTATGATGATGATGATGATGATGATG 46243
Db |||||
QY 3959 CAGATCTACTGGGGTGGACAGTACACCTGGGACATGCAAAAGCATGGCAGAGATGATGG 4018
Db |||||
QY 46244 CAGATCTACTGGGGTGGACAGTACACCTGGGACATGCAAAAGCATGGCAGAGATGATGG 46303
Db |||||
QY 4019 TGTAGTATGGATGAATTTGGAAGGGGTCTATGGTACTCAATGAGAAAGATGAGTATGAAGAT 4078
Db |||||
QY 46304 TGTAGTATGGATGAATTTGGAAGGGGTCTATGGTACTCAATGAGAAAGATGAGTATGAAGAT 46363
Db |||||
QY 4079 CAGGCCCCCTTCTCCACAGCAATAGTCCCCAATACGTAGATTTTGTCTTCTGTATGTTG 4138
Db |||||
QY 46364 CAGGCCCCCTTCTCCACAGCAATAGTCCCCAATACGTAGATTTTGTCTTCTGTATGTTG 46423
Db |||||
QY 4139 ACAACATTTTGTATCAATTTATTTGGAATTTTCTTTCATACATTAATTTCTCTCTTAAA 4198
Db |||||
QY 46424 ACAACATTTTGTATCAATTTATTTGGAATTTTCTTTCATACATTAATTTCTCTCTTAAA 46483
Db |||||
QY 4199 CTCTCAAGCAGACGCTGAGTGTGACTTTTGTAAAAAAGTATAGGATAAATTTACATTTAAAT 4258
Db |||||
QY 46484 CTCTCAAGCAGACGCTGAGTGTGACTTTTGTAAAAAAGTATAGGATAAATTTACATTTAAAT 46543
Db |||||
QY 4259 AGCAGATGATTTTCTTTTGTGTTTCTCAATTTCTTGTCTCAACCCAGAAAGTAAACAAAGT 4318
Db |||||
QY 46544 AGCAGATGATTTTCTTTTGTGTTTCTCAATTTCTTGTCTCAACCCAGAAAGTAAACAAAGT 46602
Db |||||
QY 4319 ATAGTTTTCAGACAGTGTGGTGTCTCATATTTTCAGTTCTTAGTTGATTCGAGAAATTTTCAA 4378
Db |||||

```
Db 46603 ATAGTTTTCACAGAGTGGTGTTCATAATTTTCAGTCTTAGTTGATTCGAGAAATTTTCAA 46662
QY 4379 ATAAGGAGAGGGGTCTTTATTCCTTGTCGTAGGAAACCATGACGGAAAGGAAAACTG 4438
Db 46663 ATAAGGAGAGGGGTCTTTATTCCTTGTCGTAGGAAACCATGACGGAAAGGAAAACTG 46722
QY 4439 ATGTTTAAAGTCCACTTTTAAACCTATATTTATTTATGTAGGACTCTGCAAGAAAACT 4498
Db 46723 ATGTTTAAAGTCCACTTTTAAACCTATATTTATTTATGTAGGACTCTGCAAGAAAACT 46782
QY 4499 TCCAAAAGATTTATTAATAACACAGACTCTGTTGCAATAAGTTAAATGTTTCTTGTTT 4558
Db 46783 TCCAAAAGATTTATTAATAACACAGACTCTGTTGCAATAAGTTAAATGTTTCTTGTTT 46842
QY 4559 TGTAATCCACACATCAATGAGTAGGCTTTCAGCTTGTAAAGGAGGAGAGCGTTTACA 4618
Db 46843 TGTAATCCACACATCAATGAGTAGGCTTTCAGCTTGTAAAGGAGGAGAGCGTTTACA 46902
QY 4619 ACCTCAATAGCTAATAAACCGTCTTGAATATTTGAAGATTTAAATCTGACTCTAGGA 4678
Db 46903 ACCTCAATAGCTAATAAACCGTCTTGAATATTTGAAGATTTAAATCTGACTCTAGGA 46962
QY 4679 CGGCGACGGTGGCTCAGCACTATAATCCACACACTTTGGAGGCTGAGCGGGCGGTCA 4738
Db 46963 CGGCGACGGTGGCTCAGCACTATAATCCACACACTTTGGAGGCTGAGCGGGCGGTCA 47022
QY 4739 AAGTGCAGGAGTTCAAGACCGCTGACCAATATGGTGAACCCCATCTCTACTAAAAAT 4798
Db 47023 AAGTGCAGGAGTTCAAGACCGCTGACCAATATGGTGAACCCCATCTCTACTAAAAAT 47082
QY 4799 ACAAAAATTAGCAGCGGTGGCGAGGTGCTGTAGGTCCAGCTAGCTGTGAGGTGG 4858
Db 47083 ACAAAAATTAGCAGCGGTGGCGAGGTGCTGTAGGTCCAGCTAGCTGTGAGGTGG 47140
QY 4859 AGATTGCATTGACCAAGATC 4879
Db 47141 AGATTGCATTGACCAAGATC 47161

RESULT 5
AF388026
LOCUS Homo sapiens fibrinogen, B beta polypeptide (FGB) gene, complete
DEFINITION
ACCESSION AF388026
VERSION AF388026.1 GI:14423574
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 11956)
AUTHORS Rieder,M.J., Carrington,D.P., Chung,M.-W., Lee,K.L., Poel,C.L.,
Yi,Q. and Nickerson,D.A.
TITLE Direct Submission
JOURNAL Submitted (05-JUN-2001) Molecular Biotechnology, University of
Washington, 1705 NE Pacific, Seattle, WA 98195, USA
COMMENT To cite this work please use: SeattleSNPs. NHLBI Program for
Genomic Applications, UW-FHCRC, Seattle, WA (URL:
http://pga.mbt.washington.edu).
FEATURES
source
1..11956
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
repeat_region
1..55
/rpt_family="L2"
/rpt_type="dispersed"
40
variation
/frequency="0.01"
/replace="G"
472
variation
```

```
/frequency="0.13"
/replace="A"
675
/frequency="0.01"
/replace="T"
749
/frequency="0.01"
/replace="G"
899
/frequency="0.14"
/replace="T"
1038
/frequency="0.13"
/replace="A"
1437
/frequency="0.13"
/replace="A"
1442
/frequency="0.01"
/replace="T"
1643
/frequency="0.18"
/replace="T"
1744
/frequency="0.13"
/replace="T"
41892..9964
/replace="FGB"
Join(1892..2013,4689..4880,5370..5553,6474..6701,
7262..7375,8062..8187,8394..8679,9299..9964)
/gene="FGB"
/product="fibrinogen, B beta polypeptide"
Join(1900..2013,4689..4880,5370..5553,6474..6701,
7262..7375,8062..8187,8394..8679,9299..9530)
/gene="FGB"
/codon_start=1
/product="fibrinogen, B beta polypeptide"
/protein_id="AAK62470.1"
/db_xref="GI:14423575"
/translation="MKRMSWSFHKLTKMKHLLLLCVFLVKSGVNDNERGFSAR
GHRPLDKREERAPSLRPPPIISGGYRARPAAKATQKVERKAPDAGGCLHAPDL
GVLPTGCOLQALQORPIRNSVDLNNVAVSQTSSSFQYMLLKDWKQKQK
QVKNENVEYSSLEKHLYIDVTNSNIPNLRLVLSILENLSKIQKLESVSA
QMEVCTPCTVSNIPVVSKECEILIRKGETSEMYLIQPDSSVKPRVYCDMTEN
GWTIVQNRQDGSDFGRKWDPIYKQFGVATNDGKNYCGLPBGYIWNLDKISQITR
MGPTELLMEEDWKDKVHAHYGFTVQNEANKYQISVKNYRGTAGNALMDGASQIMG
ENRTMTIHNGMFFSTYDRDNDGWLTSDPKQCKSDKGGGWMYNRCHAANPNRGYYWGG
QYTWDMAKHGTDDGVVMMNWKGSWYSMRKMSMKIRPFPFQQ"
2240
/gene="FGB"
/frequency="0.13"
/replace="A"
2360
/gene="FGB"
/frequency="0.01"
/replace="G"
2620
/gene="FGB"
/frequency="0.01"
/replace="C"
2623
/gene="FGB"
/frequency="0.02"
/replace="A"
2867
/gene="FGB"
/frequency="0.14"
/replace="A"
2933..2934
/gene="FGB"
/frequency="0.01"
/replace="A"
3395..3471
repeat_region
```


[illegible]

QY 297 TTCCGAGCTGACATCAGCAAGTGTGATTTCTTATGAAAAATTTCTACTATGACTCTTAT 356
Db 56212 TTCCGAGCTGACATCAGCAAGTGTGATTTCTTATGAAAAATTTCTACTATGACTCTTAT 56153
QY 357 TTTAAGTATACAAGAACTTGTGACTCAGAGATAATTTTACAGAGTGGAAAAAACC 416
Db 56152 TTTAAGTATACAAGAACTTGTGACTCAGAGATAATTTTACAGAGTGGAAAAAACC 56093
QY 417 CTAGCAATTAATAGTTTAAACATTTT-GAGGTTTGAATGAG-AGAGTTATCCATATATAT 474
Db 56092 CTAGCAATTAATAGTTTAAACATTTTGAATGAGGTTTGAATGAGNAGAGTTTATCCATATATAT 56033
QY 475 TCAATTTGTTGTGTGATAATGACACCTAAACC-TGTGAATCTTTGAGTTCAGAA-TGTTGAG 532
Db 56032 TCNATTTGTTGTGATAATGACACCTAACCNVTGTAATCTTTGAGTTCAGAAATGTTGAG 55973
QY 533 TGCTGTTGACTTGTGTGTGAGAAACAGCTAGTGTGAGCTGGGCAAGGCAATCTCAGT 592
Db 55972 TGCTGTTGACTTGTGTGTGAGAAACAGCTAGTGTGAGCTGGGCAAGGCAATCTCAGT 55913
QY 593 GAGTAGCATACCCACAGTTGGAAATTTTTCAGAGAAATCAAAGGAATCATGACATCTTAT 652
Db 55912 GAGTAGCATACCCACAGTTGGAAATTTTTCAGAGAAATCAAAGGAATCATGACATCTTAT 55853
QY 653 AAATTTCAAGGTTCTGCTATACCTATGTGAATGATAAATAAATCAAAGCATATCCACTC 712
Db 55852 AAATTTCAAGGTTCTGCTATACCTATGTGAATGATAAATAAATCAAAGCATATCCACTC 55793
QY 713 TGTAAGATTGAATCTCTCAGATGGAAGACCCCAATCTGCTTTCTCTTTTCCCTCAC 772
Db 55792 TGTAAGATTGAATCTCTCAGATGGAAGACCCCAATCTGCTTTCTCTTTTCCCTCAC 55733
QY 773 CAAGAAATTAACACCTATTTCAATTTATCTGACACCAATCTTTAGCGTATACCTATG 832
Db 55732 CAAGAAATTAACACCTATTTCAATTTATCTGACACCAATCTTTAGCGTATACCTATG 55673
QY 833 GTAAATTTACTAGTAGTGTGTGTAGGATTTATGTTAATTTGTATATGTCTATGCGCAAAATC 892
Db 55672 GTAAATTTACTAGTAGTGTGTGTAGGATTTATGTTAATTTGTATATGTCTATGCGCAAAATC 55613
QY 893 ATTTCCACATAATGACTAT 952
Db 55612 ATTTCCACATAATGACTAT 55553
QY 953 TTATTTCTAGAAATCAAAATTTGTATGTTAATATATATATATATATATATATATATATAT 1012
Db 55552 TTATTTCTAGAAATCAAAATTTGTATGTTAATATATATATATATATATATATATATATAT 55493
QY 1013 TAACTATTTCTACATAATTTCAATTTTCCAGATAATGAAATTTGTATGTAATGAGTACTCC 1072
Db 55492 TAACTATTTCTACATAATTTCAATTTTCCAGATAATGAAATTTGTATGTAATGAGTACTCC 55433
QY 1073 TCAGAACTGGAAAGACCAAT 1132
Db 55432 TCAGAACTGGAAAGACCAAT 55373
QY 1133 AACCTTGTGTGCTTGTGTTCAATCTTGGAAACCTTGAGAGCAAAATACAAAAGTTAGAA 1192
Db 55372 AACCTTGTGTGCTTGTGTTCAATCTTGGAAACCTTGAGAGCAAAATACAAAAGTTAGAA 55313
QY 1193 TCTGATGCTCAGCTCAAAATGGAATATTTGTCACCCCAATGCACTGTGCTGATTTGCAATATT 1252
Db 55312 TCTGATGCTCAGCTCAAAATGGAATATTTGTCACCCCAATGCACTGTGCTGATTTGCAATATT 55253
QY 1253 CCTGTGCTGCTGGCAAGGTAATCTGATTTCAATAACATATTTTATAGAGGTTCCAGAGA 1312
Db 55252 CCTGTGCTGCTGGCAAGGTAATCTGATTTCAATAACATATTTTATAGAGGTTCCAGAGA 55193
QY 1313 ACTCACACCAAAATTAAGAGAACCAACCAACCAAAATTTCTAAGTGGATTTTCCC 1372
Db 55192 ACTCACACCAAAATTAAGAGAACCAACCAACCAAAATTTCTAAGTGGATTTTCCC 55133
QY 1373 AACAGATCATATGACATACATATATATATATATATATATATATATATATATATATAT 1432

Db 55132 AACAGATCATATGACATATACATATATATATATATATATATATATATATATATATATAT 55073
QY 1433 ACTGCCCTGGTGCATTTGCTGCTTTGATGACAGGATGGGCGACAGGTAGTCCCAGGG 1492
Db 55072 ACTGCCCTGGTGCATTTGCTGCTTTGATGACAGGATGGGCGACAGGTAGTCCCAGGG 55013
QY 1493 TGGCTGATGCTGTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1552
Db 55012 TGGCTGATGCTGTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 54953
QY 1553 GAACTTTCTCATTTTGTGCTTCTGAGGCTCATGATATTTCCAGGACACAAAGGTTGG 1612
Db 54952 GAACTTTCTCATTTTGTGCTTCTGAGGCTCATGATATTTCCAGGACACAAAGGTTGG 54893
QY 1613 AGAAGAGCTTTGTTGCTCTCTTACAGATAAAGCTCTCAAACTGCGGTTGGACTTACT 1672
Db 54892 AGAAGAGCTTTGTTGCTCTCTTACAGATAAAGCTCTCAAACTGCGGTTGGACTTACT 54833
QY 1673 AAAAGTAAATGAAATCTAATTTTGTATATATATTTTCAAAGGCTCTATAAACAACACT 1732
Db 54832 AAAAGTAAATGAAATCTAATTTTGTATATATTTTCAAAGGCTCTATAAACAACACT 54773
QY 1733 CCTTAGTAATCTTATGTAATGTTTAAAGAAATTTGGTGAATTAATTAATTAATTAATTA 1792
Db 54772 CCTTAGTAATCTTATGTAATGTTTAAAGAAATTTGGTGAATTAATTAATTAATTAATTA 54713
QY 1793 TCATAAACCCTGTAACATAATGTTGCTTCAATTTGCAAGATGTGAGAAATTTATCAGGA 1852
Db 54712 TCATAAACCCTGTAACATAATGTTGCTTCAATTTGCAAGATGTGAGAAATTTATCAGGA 54653
QY 1853 AAGGAGGTGAAACATCTGAAATGTATCTCAATTCACCTGACAGATTTCTGTCAAAACGTATA 1912
Db 54652 AAGGAGGTGAAACATCTGAAATGTATTTTTCACCTGACAGATTTCTGTCAAAACGTATA 54593
QY 1913 GAGTATATGTCATGACATGAAATGAGGTAAGCTTTTCGACAGTTGTTGACCTGT 1972
Db 54592 GAGTATATGTCATGACATGAAATGAGGTAAGCTTTTCGACAGTTGTTGACCTGT 54533
QY 1973 TGATCTGTAAATTTATTTGATACCGTAAATGCGAGAAACAGGCGAGGTGGTGGCTC 2032
Db 54532 TGATCTGTAAATTTATTTGATACCGTAAATGCGAGAAACAGGCGAGGTGGTGGCTC 54473
QY 2033 ATACCTGTAAATTTCCAGCACCTTTGGAGGCAAAAGTGGCTGATAGCTTGAAGCTAGGAGT 2092
Db 54472 ATACCTGTAAATTTCCAGCACCTTTGGAGGCAAAAGTGGCTGATAGCTTGAAGCTAGGAGT 54413
QY 2093 TTGAAACTAGCTGGGCAACATAATGAGACCCCTACTCTCAAAAACCAAAATATACCA 2152
Db 54412 TTGAAACTAGCTGGGCAACATAATGAGACCCCTACTCTCAAAAACCAAAATATACCA 54354
QY 2153 AAAAAACCAAAATATGAGTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 2212
Db 54353 AAAAAACCAAAATATGAGTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 54294
QY 2213 -GGCTGAGAT-GGGAGATCACCTGAGCCCAACCT-GGAGTCTTGTATCATGCTACTGAA 2269
Db 54293 GGGCTGAGATGGGAGATCACCTGAGCCCAACCTGGAGGCTTGTATCATGCTACTGAA 54234
QY 2270 CTGTAGCTGGCAACAGAGGATGAGATCTGCTCTCAAAAACCAAAATATTAATTAATA 2329
Db 54233 ACTGAACCTGGCAACAGAGGATGAGATCTGCTCTCAAAAACCAAAATATTAATTAATA 54174
QY 2330 AGCCAGGAAACAAAGCTTAGCTTAAACATCTAACATAGCTGACAAA--ATAATTTGATG 2389
Db 54173 AGCCAGGAAACAAAGCTTAGCTTAAACATCTAACATAGCTGACAAA--ATAATTTGATG 54116
QY 2390 TGAATTTCAACCTGATATTTAAAGTTATAAATATCTATATATTTCAAAATTTGGGTAAG 2449
Db 54115 TGAATTTCAACCTGATATTTAAAGTTATAAAGTTATAAATATCTTAAATTTCAAAATTTGGGTAAG 54056
QY 2450 ATAAAGCTTTCAGTTTCAAGATTTTACAGTTTACCTCTCATATTTTCTCTTAT 2509

Qy	4789	TACTATAAAATACAAAAATTAGCCAGGCGTGGTGCGCAGGTGCTGTAGGTGCCAGCTAGCC	4848
Db	49158	TACTATAAAATACAAAAATTAGCCAGGCGTGGTGCGCAGGTGCTGTGTA-GTCCAGCTA-CC	49215
Qy	4849	TGTGAGGTGAGATTGCATTGAGCCCAAGATC	4879
Db	49216	TGTGAGGTGAGATTGCATTGAGCCCAAGATC	49246
RESULT 8			
LOCUS	AX535001		
DEFINITION	Sequence 28 from Patent WO02068633.	4043 bp	DNA
ACCESSION	AX535001		linear
VERSION	AX535001.1	GI:25261604	PAT 22-NOV-2002
KEYWORDS			
SOURCE			
ORGANISM	Homo sapiens (human)		
REFERENCE			
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
TITLE	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
JOURNAL	1		
FEATURES			
source			
ORIGIN			
Query Match	42.7%;	Score 2081.2;	DB 6; Length 4043;
Best Local Similarity	99.9%;	Pred. No. 0;	
Matches 2083;	Conservative	0; Mismatches	3; Indels 0; Gaps 0;
Qy	2018	CAGGTGTGGTGCTCATCTGTAATTCAGCACCTTGGGAGGCCAAAGTGGGTGTATAG	2077
Db	1958	CAAGTGTGGTGCTCACCTGTAATCCAGCACCTTGGGAGGCCAAAGTGGGTGTATAG	2017
Qy	2078	CTTGAGCCTAGGAGTTGAAACTAGCTGGGCAACATAATGAGACCCCTAACTCTACAAA	2137
Db	2018	CTTGAGCCTAGGAGTTGAAACTAGCTGGGCAACATAATGAGACCCCTAACTCTACAAA	2077
Qy	2138	AAAAAAAATACCAAAAAAATAATCAGCTGTGTTGGTAGTATGTCCTGTAGT	2197
Db	2078	AAAAAAAATACCAAAAAAATAATCAGCTGTGTTGGTAGTATGTCCTGTAGT	2137
Qy	2198	CCGAGCTATCCAGGAGCTGAGTGGGAGATCACTGAGGCCCAACCTGAGGTCTTGAT	2257
Db	2138	CCGAGCTATCCAGGAGCTGAGTGGGAGATCACTGAGGCCCAACCTGAGGTCTTGAT	2197
Qy	2258	CATGCTACTGAACTGTAGCCTGGGCAACAGAGCATAGTGAGATCTGTCTCAAAAAAAA	2317
Db	2198	CATGCTACTGAACTGTAGCCTGGGCAACAGAGCATAGTGAGATCTGTCTCAAAAAAAA	2257
Qy	2318	AATTAAATTAAGGACAGAAACAAGACTTGTAGCTCTAACTTAACATGAGCTGCAAGG	2377
Db	2258	AATTAAATTAAGGACAGAAACAAGACTTGTAGCTCTAACTTAACATGAGCTGCAAGG	2317
Qy	2378	AGTAATTTGATGTGGAATTCACCTGATATTTAAAGTTATAAATATCTATANTTCACA	2437
Db	2318	AGTAATTTGATGTGGAATTCACCTGATATTTAAAGTTATAAATATCTATANTTCACA	2377
Qy	2438	ATTGGGGTGAAGATAAGCACTTGAGAGTTTCCAAAGATTTTACAAGTTTACCTCTCATAT	2497
Db	2378	ATTGGGGTGAAGATAAGCACTTGAGAGTTTCCAAAGATTTTACAAGTTTACCTCTCATAT	2437
Qy	2498	TTATTTCCCTATTGTGCTATTTTAGAGCAACAAATATATATACTAAATGGAATGACAGG	2557
Db	2438	TTATTTCCCTATTGTGCTATTTTAGAGCAACAAATATATATACTAAATGGAATGACAGG	2497

Millennium Predictive Medicine, Inc. (US)

FEATURES

Location/Qualifiers
1. 4967
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

source

ORIGIN

Query Match 19.9%; Score 971.8; DB 6; Length 4967;

Best Local Similarity 99.5%; Pred. No. 4e-190; Matches 1006; Conservative 0; Mismatches 2; Indels 3; Gaps 3;

QY 3869 CAGGTTAAACATCAGATCCAGAAAACAGTGTCTAAAGAAAGACGGTGGTGGATGGTGA 3928
DB 2428 CTGGTTAAACATCAGATCCAGAAAACAGTGTCTAAAGAAAGACGGTGGTGGATGGTGA 2369
QY 3929 TAATAGATGTCTATGACGCAATCCAAACGGCAGATCTACTCGGGTGGACAGTACACCTG 3988
DB 2368 TAATAGATGTCTATGACGCAATCCAAACGGCAGATCTACTCGGGTGGACAGTACACCTG 2309
QY 3989 GGACATGCCAAGCATGGCACAGATGATGCTGTAGTATGGATGAATTCGAAGGGGTGATG 4048
DB 2308 GGACATGCCAAGCATGGCACAGATGATGCTGTAGTATGGATGAATTCGAAGGGGTGATG 2249
QY 4049 GTACTCAATGAGGAAGATGAGTATGAAGATCAGGCGCTTCTTCCACAGCAATAGTCCCC 4108
DB 2248 GTACTCAATGAGGAAGATGAGTATGAAGATCAGGCGCTTCTTCCACAGCAATAGTCCCC 2189
QY 4109 AATAGTATGATTTTGTCTTCTCTATATGACAAATTTTGTGATCATATGTTATTTGAA 4168
DB 2188 AATAGTATGATTTTGTCTTCTCTATATGACAAATTTTGTGATCATATGTTATTTGAA 2129
QY 4169 TTTTCTTTCATACATATATCTCTCTAACTCTCAAGCAGACGGTGGTGGATCTTTTGT 4228
DB 2128 TTTTCTTTCATACATATATCTCTCTAACTCTCAAGCAGACGGTGGTGGATCTTTTGT 2069
QY 4229 AAAAAGTATAGGATAAATTAATTAATAATAGCATGATTTCTTTTGTCTTTCTCAT 4288
DB 2068 AAAAAGTATAGGATAAATTAATTAATAATAGCATGATTTCTTTTGTCTTTCTCAT 2009
QY 4289 TCTTTGTCTACCCAGAAAGTAAACAAAGTATAGTTTGCAGAGATGGTGGTTCATAAT 4348
DB 2008 TCTTTGTCTCA-CCAAAGAAAGTAAACAAAGTATAGTTTGCAGAGATGGTGGTTCATAAT 1950
QY 4349 TCAGTCTAGTTGATGTCGAGATTTTCAAAATAGGAAGAGGGTCTTTTATCCTTGTG 4408
DB 1949 TCAGTCTAGTTGATGTCGAGATTTTCAAAATAGGAAGAGGGTCTTTTATCCTTGTG 1890
QY 4409 TAGGAAAACCATGACGGAAGAAAACCTGATGTTTAAAGTCCACTTTTAAACCTATAT 4468
DB 1889 TAGGAAAACCATGACGGAAGAAAACCTGATGTTTAAAGTCCACTTTTAAACCTATAT 1830
QY 4469 TTATTTATGATGATCTGTCAAGAAAACCTTCAAAAGATTTATTAATAAACCCAGAT 4528
DB 1829 TTATTTATGATGATCTGTCAAGAAAACCTTCAAAAGATTTATTAATAAACCCAGAT 1770
QY 4529 CTGTTGCAATAGTTAAATGTTTCTTGTCTTGTATTCACATCAATGAGTGGCTT 4588
DB 1769 CTGTTGCAATAGTTAAATGTTTCTTGTCTTGTATTCACATCAATGAGTGGCTT 1710
QY 4589 TGCACCTTTGAAGGAAGGAAGCGTTTCAAACTCAATAGCTAATAAACCCGCTCTTGA 4648
DB 1709 TGCACCTTTGAAGGAAGGAAGCGTTTCAAACTCAATAGCTAATAAACCCGCTCTTGA 1650
QY 4649 TATTTGAAGATTTAAATCTGACTCTAGACGGGCAAGCGTGGCTCACGACTAATAATCCA 4708
DB 1649 TATTTGAAGATTTAAATCTGACTCTAGACGGGCAAGCGTGGCTCACGACTAATAATCCA 1590
QY 4709 ACATTTGGGAGCTGAGCGGGCGGTCTACAGGTCAGAGTTCAAGACAGCTTGACCA 4768
DB 1589 ACATTTGGGAGCGGAGCGGGCGGTCTACAGGTCAGAGTTCAAGACAGCTTGACCA 1530
QY 4769 ATATGTTGAAACCCCATCTCTCTATAAAATACAAAAATTAGCCAGGCGTGGTGGCAGGTG 4828

DB 1529 ATATGTTGAAACCCCATCTCTACTAAAAATACAAAATTAGCCAGCGTGGTGGCAGGTG 1470
QY 4829 CCTGTAGTCCAGCTAGCTAGCTGTGAGTGGAGATTCATTGAGCCCAAGATC 4879
DB 1469 CCTGTA-GTCCACAGCTA-CCTGTGAGTGGAGATTCATTGAGCCCAAGATC 1421

RESULT 11

AC140369/c

LOCUS

DEFINITION

AC140369 Mus musculus chromosome UNK clone RP23-200L3, WORKING DRAFT

SEQUENCE, 9 unordered pieces.

ACCESSION

AC140369

VERSION

AC140369.1 GI:28475627

KEYWORDS

HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.

SOURCE

Mus musculus

ORGANISM

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 223534)

McPherson, J.D. and Waterston, R.H.

TITLE

The sequence of Mus musculus clone

Unpublished

REFERENCE

2 (bases 1 to 223534)

McPherson, J.D. and Waterston, R.H.

Direct Submission

Submitted (23-FEB-2003)

Genome Sequencing Center, 4444 Forest Park

Parkway, St. Louis, MO 63108, USA

COMMENT

----- Genome Center -----

Center: Washington University Genome Sequencing Center

Center code: WUGSC

Web site: http://genome.wustl.edu/gsc/index.shtml

Contact: submissions@watson.wustl.edu

----- Project Information -----

Center project name: M.BA020103

----- Summary Statistics -----

Sequencing vector: M13; 0%

Chemistry: Dye-primer ET; 0% of reads

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.990319

Consensus quality: 218876 bases at least Q40

Consensus quality: 219375 bases at least Q30

Consensus quality: 219567 bases at least Q20

----- NOTE: This is a 'working draft' sequence. It currently

* consists of 9 contigs. The true order of the pieces

* is not known and their order in this sequence record is

* arbitrary. Gaps between the contigs are represented as

* runs of N, but the exact sizes of the gaps are unknown.

* This record will be updated with the finished sequence

* as soon as it is available and the accession number will

* be preserved.

* 1

* 4212: contig of 4212 bp in length

* 4312: gap of unknown length

* 4313: contig of 12516 bp in length

* 16928: gap of unknown length

* 16929: contig of 12577 bp in length

* 29506: gap of unknown length

* 29506: gap of unknown length

* 42552: contig of 12947 bp in length

* 42553: gap of unknown length

* 59543: contig of 16891 bp in length

* 59544: gap of unknown length

* 59644: contig of 30976 bp in length

* 90620: gap of unknown length

* 90720: contig of 23418 bp in length

* 11438: gap of unknown length

* 11438: contig of 45164 bp in length

* 159402: gap of unknown length

* 159502: contig of 64033 bp in length.

1688 TCATAATTTGTTATATATTTTCAAAGTCTATATATACACACACCTCTTAGTAACCTTAG 1747
167491 TCCAACTCTAGAAACATTAATGTTAGATGTTTGTGATATAAATGGTCTCTTACCAGTGG 167432
1748 TAAATGTTATTTAAAGAAATGGTGACTATAATAAAGTAATATATGTCATPAAACCCCTGAA 1807
167431 AACAAATCAGTTATGTTCAAGAGGAATAGCTACA---GGATCTGTTACTGTCCCTGAACA 167375
1808 CATATATGTTGTTTACATTTGAGAAATGTCAGAAATATATCAGGAAGAGGAGTGAACAT 1867
167374 TAAATGTTGTCATCATATTTCTGAGAGTGTGAGGAGATCATATAGGAAGGAGGTGAGACAT 167315
1868 CTGAATATGTTCTCATTTCAACCTGACAGTCTGTGCAACCGTATAGAGTATAGTGTGACA 1927
167314 CCGAAATGTTATCTATCCAGCTGACACCTCCATCAGCCGTCACAGATATAGTGTGACA 167255
1928 TGAATACAGAAATGAGGTAGCTTTTCGACAGTGTGTTGACCTGTTGATCTGTAATATT 1987
167254 TGAATAACAGAAACGAGGTAGTGTGATGCTCAGTGGCTGACCTGTCATTTCTGTAAGCATC 167195
1988 TGGATACCGTAAATGCGCAGGAAACAGGCCAGGTGTGGTGGCTCATACCTGTAATTTCCA 2047
167194 CTGGGTGCTGGGGAACAAGG-----167174
2048 GCACCTTGGGAGCCCAAGTGGGCTGATAGCTTTGAGCTAGGAGTTTGAACCTAGCCTGG 2107
167173 -----167174
2108 GCAACATATGAGACCTTAACCTTACAAAAAATAAATAAATAAATAAATAAATAAATAA 2167
167173 -----167174
2168 TCAGCTGTGTTGGTAGTAGTGTGCTGTAGTCCAGCTATCCAGGAGGCTGAGATGGAGA 2227
167173 -----167174
2228 TCACCTGAGCCCAACCTGGAGTCTTTGATCATGCTGTAAGTGTGAGCTGGGCAACAG 2287
167173 -----167174
2288 AGGATAGTGAGATCTGTCTCAAAAAAATAAATAAATAAATAAATAAATAAATAAATAA 2347
167173 -----167174
2348 AGCTCTAACATCTAACATAGCTGACAAAGAGTAATTTGATGTGGTAATTCACCTGATAT 2407
167173 -----TTACGCTCCAAATTCAGATGACAAGAACCACTT-----167141
2408 TTAAGATTTATAAATATCTATAATTCACAAATTTGGGGTAAGATAAAGCACTTGCAATTT 2467
167140 -----GATTTGAAGAACTCGATGCGATGATGAAGGTGTGATGACAGCACTTGAGATTT 167087
2468 CCAAGATTTTCAAGTTTACCTCTCATATTTATTTCTTATTTGTTCTATTTT-----TAG 2523
167086 CCAATATCTTTAAGAGTCTCTCTTTATGCTTATTTCTTATTCATCATCTCCCTTGGAGAG 167027
2524 AGCACCATAATATCTAAATGGAATGGACAGGGATTCAGATATATTTTCAAAGTGAC 2583
167026 GGACCATAGAGGCAATAGATAAATAAATAAATTTGGCAATAGTTTCAGACTTTTCAGAGTAAT 166967
2584 ATTATTTGCTGTTGTTATATATGCTCTTTTGTCTCTGTCACACCAAGAGATGGACAGT 2643
166966 GGTATTTAG-----TGATTTCTTAGGCTAAATTTTATTTCTCTCAACTAAAGGATGGACGT 166911
2644 GATTTCAAGACCGTCAAGACGGTAGTGTGACATTTGGCAGGAATGGGATCCCATATAAACA 2703
166910 CATACAGAACCGTCAAGATGGACGCTCGACITTTGGCAGGAATGGGACCCATACAAAA 166851
2704 GGGATTTGAAATGTTGCAACCAACACAGATGGGAAGAAATTTACTGTGGCTTACCAGGTAA 2763
166850 AGGATTCGGAATATTTGCCCAACCAATGAAGACGGAAGAGTACTGTGGCTTGCAGGTAA 166791

2764 CGAACAGCGATGCACAAATAAATAATCATTTATTTGAAATGGGATTTTTTTTAAATTAATAA 2823
166790 AGACTGCGAGTGAAAAAATCAATCCTTCTATTTTAAAGAGTTTTTTTCTCAGTCAAAA 166731
2824 CATTCATTTGTTGAAGCGCTGTTTTAGG-----2850
166730 CATCAAGGATGGGATCTCTTTTTAGAAATTTTAGGATGCCAGGAAGAAATTTAACTACAC 166671
2851 -----CAGTTAAGAGGAGTTTCTCGACAAAAATGTCGACAAAAATGTCGAGCT 2887
166670 ACATGCACTACACACATATGCACATAAATAAATGTAATGTTTAAATGTCGAGCT 166611
2888 AAGATAGGGAAGAAAGCAGTTTTTTAGTTTCCCAAAATTTTATTTTGGTGAGAGATT 2947
166610 AAGGATCTGGAGGAA--GCTGTGCCCGTTGCCCTCGCTTTACTTTTGTGATGGCTT 166553
2948 TTAATTTGTTTTCTTTTAGGTGAATATTTGGTTCGAAATGATAAATTTAGCAGCTTAC 3007
166552 ATA---TCCATTTCTGTAGGTGAATATTTGGCTGGCAATGATAAGATTAGCCAGCTCAC 166496
3008 CAGGATGGGACCCACAGAACTTTTGTAGTAATGAGGAGCTGGAAGGAGACAAAGTAAA 3067
166495 CAGGATGGGACCCACAGAACTTTCTCATTTGAATGAGGAGCTGGAAGGAGACAAAGTAAA 166436
3068 GGCTCACTATGAGGATTCACCTGTACAGAAATGCAAGCAAAATACCAAGATCTCAGTGAA 3127
166435 GGCACTATGAGGCTTTCAGGTACAGAAAGGAGGCGCAATATACCAAGTCTCAGTGAA 166376
3128 CAAATACAGAGAAACAGCGGTAAATGCTCATGATGAGGAGCTCTCAGCTGATGGAGA 3187
166375 CAAATACAAAGGACGCGGTGCAACGCTCATGACGCGGCTCTCAACTGGTGGGGA 166316
3188 AAGAGGACCATGACCATTCACAAAGGAGTGTCTTTCAGCAGTATGACAGACATGCA 3247
166315 GAACAGAACCATGACCATTCACAAAGGAGTGTCTTTCAGCAGTATGACAGGAGCAACGA 166256
3248 CGGCTGGTATGTTGGCACTTTTGTCTCTG--CTTTAAATAACACATAATATCACTACT 3306
166255 CGGCTGGTATGACCAAGGTCTTAGCTCTGCTTGTGATGATGACCCCAAGATGGATTGTC 166196
3307 CAGAAATCAATA-----CAATATTTTAAATAGTACCACTTCTCTGGGCACTTA-----3354
166195 TAGCACTGTGGGGTACCGAATGTTCTAAACAACCTGCCAGACTTGGACACTTCAAGTCT 166136
3355 --CTGTGAGCACTGCTCTAGCTCTTTATGATCACTCTGAAAGCATTTCAACTATAAGG 3412
166135 ATCAGGACGCCCCCAATCCGAACTCTTACGTAATCAAGGGAAGCATCCCGCTTAAGA 166076
3413 TAGCAATCTTATTTCTCATTTTACAGATGAGATTTAGAGAGATTTAGCTGATTTGTCAT 3472
166075 CAGACATGTTAGTTTCTCTTTTCTAGATGAGTTTATGATACGATTTTACCC-----166023
3473 GTTCACAACTACCCAGAGATAAATACTAGAAATTTGAGCAGCTTACTTTCTGAATAAGA 3532
166022 --TAGACAGCTACCCAAAGATAGACTGTATCTATAGCACAAAC-----TCTGAGTTTTGT 165969
3533 GCATTTAGATAAATACCTATATCTCTATTTCTTAAAGTGTGTGTGAAACATTTTCAATTC 3592
165968 GTCTCTATATACATCTGTGTTTTTCACTCAACGCTATCTACGGAAGCCATCATTTCCA 165909
3593 ATTTCCAGGTTCTCTGATCTAAGGTTGTAAGAGCTATTTCCAGTATATAAGTAAACA 3652
165908 TATCTCAGGCT-----CTCATCTACGATATCAGTATAAG 165872
3653 AACACAGCTCCCTAGATGAGTTG--CCAAAAAGGCCAGTTATCTCTCTTTCTGCTATAGG 3711
165871 AATTCAGTCCCTTCATGATCAGATCAGATAGACCCAGCTGTTTCCCTTTTCATGCTATGG 165812
3712 GCACAGGAGTCTTTGGTGTATTTAGTGTGACTATATGATATAGCACCACAAAGGAAGACTA 3771
165811 GCAGAGAAGAACTTGGCTGTATTTG---GGCTTTACATACAGGATTCAGGGAAGAGGT 165755
3772 CTGTGCACAGAGTGTAGCAGTCTTTTATGGGTAACTCTGCAAAACGTAACCTTTGACCACCG 3831


```

165754 CCACACACAGTGGAGTACAATACCTTTTCATCGGGGTACTCTCCAAAGGACTGAATTGATGG 165695
3832 TAGTCTCTGTTCTTAATACGCCAACA--CATTTCTTTTCAGGTTAAACATCAGATCCAG 3889
165694 TGGCAGCTGTATGATGATCACTAATCACTACATCATCTCTCTTTTCAGGGTAACTACGATCCAG 165635
3890 AAAACAGTGTCTTAAAGAGCGGTGGTGGATGGTGTATTAATAGATCTCATGCGAGCCAA 3949
165634 AAAACAGTGTCTTAAAGAGCGGTGGTGGATGGTGTATTAATAGATCTCATGCGAGCCAA 165575
3950 TCCAAACCGGAGATCACTACTCGGTGGGAGCATGACACCTGGGACATGGCAAGCATGGGAC 4009
165574 TCCAAATGGCAGATCACTACTCGGTGGGAGCATGACACCTGGGAGCATGGCAAGCATGGGAC 165515
4010 AGATGATGCTGTAGTATGATGATGAATTTGAAGGGGTGATGTTACTCAATGAGAGATGAG 4069
165514 AGATGATGCTGTAGTATGATGATGAATTTGAAGGGGTGATGTTACTCAATGAGAGATGATC 165455
4070 TATGAAGATCAGGCGCTCTTCCACACAGCAATAGTCCCAATACGATGATTTTGCCTCTT 4129
165454 TATGAAGATCAGGCGCTCTTCCACACAGCAATAGTCCCAATACGATGATTTTGCCTCTT 165395
4130 CTGATGTGCAACAATTTTGTATGATGATTTTGTGGAATTTCTTTCATACATATATTT 4189
165394 TTCTATGT-CTGACATTTTTCATATGATTTTGTGGAATTTCTTTCATACATATAT-T 165337
4190 CCTCTAAACCTCAAGCAGAGCTGAGTGTGCTCTTTTGAAGAGTATAGATAAATTA 4249
165336 CCACTAAAGCTCAAAACATGTTGAGTGTGCTCTTTTGGAACTATAGATAAAGAC 165277
4250 CATTAAATAGACATGATTTTCTTTTGTGTTCTTTCATTTCTC 4292
165276 ATTGAAGACAGCACCTAGATTTCTTCTGTTATTTCTTTCGTTTC 165234

RESULT 12
AC138394/c 227600 bp DNA linear ROD 13-APR-2004
DEFINITION Mus musculus chromosome 3, clone RP23-333D8, complete sequence.
ACCESSION AC138394
VERSION AC138394.7 GI:46359988
KEYWORDS HTG.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
1 (bases 1 to 227600)
Birnren,B., Nusbbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S.,
Barna,N., Bastien,V., Bloom,T., Boguslavskiy,L., Boukhgalter,B.,
Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A.,
Cook,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
Faro,S., Ferreira,P., FitzGerald,M., Gage,D., Galagan,J., Gardyna,S.,
Hafez,N., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R.,
Lindblad-Toh,K., Liu,X., Lui,A., Mabbitt,R., Maclean,C.,
Macdonald,P., Major,J., Manning,J., Matthews,C., McCarthy,M.,
Melidrim,J., Meneus,L., Mihova,T., Mienga,V., Murphy,T., Naylor,J.,
Nguyen,C., Nicol,R., Norbu,C., O'Connor,T., O'Donnell,P., O'Neill,D.,
Oliver,J., Peterson,K., Rachevka,A., Ramasamy,U., Raymond,C.,
Retta,R., Rise,C., Rogov,P., Roman,J., Schauer,S., Schupback,R.,
Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N.,
Stojanovic,N., Talamas,J., Testaye,S., Theodore,J., Topham,K.,
Travers,M., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X.,
Wyman,D., Young,G., Zainoun,J., Zembek,L., Zainoun,J., Zembek,L.,
Zimmer,A. and Zody,M.
Submitted (28-DEC-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA

REFERENCE
1 (bases 1 to 227600)
Birnren,B., Nusbbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S.,
Barna,N., Bastien,V., Bloom,T., Boguslavskiy,L., Boukhgalter,B.,
Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A.,
Cook,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
Faro,S., Ferreira,P., FitzGerald,M., Gage,D., Galagan,J., Gardyna,S.,
Hafez,N., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R.,
Lindblad-Toh,K., Liu,X., Lui,A., Mabbitt,R., Maclean,C.,
Macdonald,P., Major,J., Manning,J., Matthews,C., McCarthy,M.,
Melidrim,J., Meneus,L., Mihova,T., Mienga,V., Murphy,T., Naylor,J.,
Nguyen,C., Nicol,R., Norbu,C., O'Connor,T., O'Donnell,P., O'Neill,D.,
Oliver,J., Peterson,K., Rachevka,A., Ramasamy,U., Raymond,C.,
Retta,R., Rise,C., Rogov,P., Roman,J., Schauer,S., Schupback,R.,
Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N.,
Stojanovic,N., Talamas,J., Testaye,S., Theodore,J., Topham,K.,
Travers,M., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X.,
Wyman,D., Young,G., Zainoun,J., Zembek,L., Zainoun,J., Zembek,L.,
Zimmer,A. and Zody,M.
Submitted (28-DEC-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA

```

REFERENCE

AUTHORS

JOURNAL

TITLE

JOURNAL

REFERENCE

AUTHORS

JOURNAL

TITLE

JOURNAL

REFERENCE

AUTHORS

JOURNAL

TITLE

JOURNAL

REFERENCE

AUTHORS

JOURNAL

TITLE

JOURNAL

REFERENCE

AUTHORS

JOURNAL

TITLE

JOURNAL

REFERENCE

AUTHORS

JOURNAL

TITLE

JOURNAL

REFERENCE

AUTHORS

JOURNAL

TITLE

JOURNAL

REFERENCE

AUTHORS

JOURNAL

TITLE

JOURNAL

REFERENCE

AUTHORS

JOURNAL

TITLE

JOURNAL

REFERENCE

AUTHORS

JOURNAL

TITLE

JOURNAL

REFERENCE

AUTHORS

JOURNAL

TITLE

JOURNAL

REFERENCE

AUTHORS

JOURNAL

TITLE

JOURNAL

REFERENCE

AUTHORS

JOURNAL

TITLE

JOURNAL

REFERENCE

AUTHORS

JOURNAL

TITLE

JOURNAL

REFERENCE

AUTHORS

JOURNAL

TITLE

JOURNAL

REFERENCE

AUTHORS

JOURNAL

TITLE

JOURNAL

REFERENCE

AUTHORS

JOURNAL

TITLE

JOURNAL

REFERENCE

AUTHORS

JOURNAL

TITLE

JOURNAL

REFERENCE

AUTHORS

JOURNAL

TITLE

JOURNAL

REFERENCE

AUTHORS

JOURNAL

TITLE

JOURNAL

REFERENCE

AUTHORS

JOURNAL

TITLE

JOURNAL

REFERENCE

AUTHORS

JOURNAL

TITLE

JOURNAL

REFERENCE

AUTHORS

JOURNAL

TITLE

JOURNAL

REFERENCE

AUTHORS

JOURNAL

TITLE

JOURNAL

REFERENCE

AUTHORS

JOURNAL

TITLE

JOURNAL

REFERENCE

AUTHORS

JOURNAL

TITLE

JOURNAL

REFERENCE

AUTHORS

JOURNAL

TITLE

JOURNAL

REFERENCE

AUTHORS

JOURNAL

TITLE

JOURNAL

REFERENCE

AUTHORS

JOURNAL

TITLE

JOURNAL

REFERENCE

AUTHORS

JOURNAL

TITLE

JOURNAL

REFERENCE

AUTHORS

JOURNAL

TITLE

JOURNAL

REFERENCE

AUTHORS

JOURNAL

TITLE

JOURNAL

REFERENCE

AUTHORS

JOURNAL

TITLE

JOURNAL

REFERENCE

AUTHORS

JOURNAL

TITLE

JOURNAL

REFERENCE

AUTHORS

JOURNAL

TITLE

JOURNAL

REFERENCE

AUTHORS

JOURNAL

TITLE

JOURNAL

REFERENCE

AUTHORS

JOURNAL

TITLE

JOURNAL

REFERENCE

AUTHORS

JOURNAL

TITLE

JOURNAL

REFERENCE

AUTHORS

JOURNAL

TITLE

JOURNAL

REFERENCE

AUTHORS

JOURNAL

TITLE

JOURNAL

REFERENCE

AUTHORS

JOURNAL

TITLE

JOURNAL

REFERENCE

AUTHORS

JOURNAL

TITLE

JOURNAL

REFERENCE

AUTHORS

JOURNAL

TITLE

JOURNAL

REFERENCE

AUTHORS

JOURNAL

TITLE

JOURNAL

REFERENCE

AUTHORS

JOURNAL

TITLE

JOURNAL

REFERENCE

AUTHORS

JOURNAL

TITLE

JOURNAL

REFERENCE

AUTHORS

JOURNAL

TITLE

JOURNAL

REFERENCE

AUTHORS

JOURNAL

TITLE

JOURNAL

REFERENCE

AUTHORS

JOURNAL

TITLE

JOURNAL

REFERENCE

AUTHORS

JOURNAL

TITLE

JOURNAL

REFERENCE

AUTHORS

JOURNAL

TITLE

JOURNAL

REFERENCE

AUTHORS

JOURNAL

TITLE

JOURNAL

REFERENCE

AUTHORS

JOURNAL

TITLE

JOURNAL

REFERENCE

AUTHORS

JOURNAL

TITLE

JOURNAL

REFERENCE

AUTHORS

JOURNAL

TITLE

JOURNAL

REFERENCE

AUTHORS

JOURNAL

TITLE

JOURNAL

REFERENCE

AUTHORS

JOURNAL

TITLE

JOURNAL

REFERENCE

AUTHORS

JOURNAL

TITLE

JOURNAL

REFERENCE

AUTHORS

JOURNAL

TITLE

JOURNAL

REFERENCE

AUTHORS

JOURNAL

TITLE

JOURNAL

REFERENCE

AUTHORS

JOURNAL

TITLE

JOURNAL

REFERENCE

AUTHORS

JOURNAL


```
repeat_region 1. .267
/rpt_family="L1_MM"
repeat_region 269. 319
/rpt_family="A-rich"
repeat_region 349. 464
/rpt_family="PB1D7"
repeat_region 515. 658
/rpt_family="URR1A"
repeat_region complement(753. 832)
/rpt_family="BCL_MM"
repeat_region 1057. 1088
/rpt_family="CAGA)n"
repeat_region 1091. 1147
/rpt_family="GA-rich"
repeat_region complement(2211. 2449)
/rpt_family="B4A"
repeat_region 4164. 4310
/rpt_family="B1_MM"
repeat_region 4600. 4907
/rpt_family="ORR1A3"
repeat_region 5091. 5119
/rpt_family="(TTTG)n"
repeat_region complement(5120. 5266)
/rpt_family="B1_MM"
repeat_region complement(5369. 5515)
/rpt_family="MTE"
repeat_region 5627. 5857
/rpt_family="RLTR14"
repeat_region 5858. 6039
/rpt_family="(CTTG)n"
repeat_region 6040. 6164
/rpt_family="RLTR14"
repeat_region complement(6559. 7236)
/rpt_family="L1_MM"
repeat_region complement(7237. 7373)
/rpt_family="B1_MM"
repeat_region complement(7374. 7936)
/rpt_family="L1_MM"
repeat_region 7934. 9190
/rpt_family="L1"
repeat_region 10167. 10291
/rpt_family="Trigger6a"
repeat_region 10311. 10728
/rpt_family="Trigger6a"
misc_feature 10920. 10925
/Note="clone_boundary
clone_end:SP6
site:ECORI"
repeat_region 11324. 11522
/rpt_family="B3"
repeat_region 11748. 12040
/rpt_family="B4"
repeat_region 12268. 12302
/rpt_family="(TA)n"
repeat_region complement(12965. 13038)
/rpt_family="ID5"
repeat_region complement(13425. 14334)
/rpt_family="Lx9"
repeat_region complement(14508. 14803)
/rpt_family="LIM4"
repeat_region 15320. 15486
/rpt_family="RSINE1"
repeat_region 15556. 15607
/rpt_family="(CAGAGA)n"
repeat_region complement(15710. 15899)
/rpt_family="LIME3A"
repeat_region 17549. 17775
/rpt_family="B3"
repeat_region 17969. 18172
/rpt_family="B3"
repeat_region 19415. 19531
/rpt_family="B1_MM"

clone_end:SP6"
/rpt_family="L1_MM"
269. 319
/rpt_family="A-rich"
349. 464
/rpt_family="PB1D7"
515. 658
/rpt_family="URR1A"
complement(753. 832)
/rpt_family="BCL_MM"
1057. 1088
/rpt_family="CAGA)n"
1091. 1147
/rpt_family="GA-rich"
complement(2211. 2449)
/rpt_family="B4A"
4164. 4310
/rpt_family="B1_MM"
4600. 4907
/rpt_family="ORR1A3"
5091. 5119
/rpt_family="(TTTG)n"
complement(5120. 5266)
/rpt_family="B1_MM"
complement(5369. 5515)
/rpt_family="MTE"
5627. 5857
/rpt_family="RLTR14"
5858. 6039
/rpt_family="(CTTG)n"
6040. 6164
/rpt_family="RLTR14"
complement(6559. 7236)
/rpt_family="L1_MM"
complement(7237. 7373)
/rpt_family="B1_MM"
complement(7374. 7936)
/rpt_family="L1_MM"
7934. 9190
/rpt_family="L1"
10167. 10291
/rpt_family="Trigger6a"
10311. 10728
/rpt_family="Trigger6a"
10920. 10925
/Note="clone_boundary
clone_end:SP6
site:ECORI"
11324. 11522
/rpt_family="B3"
11748. 12040
/rpt_family="B4"
12268. 12302
/rpt_family="(TA)n"
complement(12965. 13038)
/rpt_family="ID5"
complement(13425. 14334)
/rpt_family="Lx9"
complement(14508. 14803)
/rpt_family="LIM4"
15320. 15486
/rpt_family="RSINE1"
15556. 15607
/rpt_family="(CAGAGA)n"
complement(15710. 15899)
/rpt_family="LIME3A"
17549. 17775
/rpt_family="B3"
17969. 18172
/rpt_family="B3"
19415. 19531
/rpt_family="B1_MM"

repeat_region 19637. 19677
/rpt_family="(TA)n"
repeat_region complement(20163. 20472)
/rpt_family="L2"
repeat_region 20521. 20558
/rpt_family="(TG)n"
repeat_region complement(20687. 21055)
/rpt_family="MT2A"
repeat_region complement(21237. 21669)
/rpt_family="L2"
repeat_region 22687. 22713
/rpt_family="(TTTG)n"
repeat_region complement(22716. 22838)
/rpt_family="B1_MM"
repeat_region 23789. 24147
/rpt_family="ORR1D"
repeat_region complement(24968. 25107)
/rpt_family="B1_MM"
repeat_region 25422. 25466

Query Match 14.3%; Score 696.2; DB 10; Length 227600;
Best Local Similarity 56.1%; Pred. No. 1.9e-133;
Matches 2480; Conservative 0; Mismatches 1298; Indels 645; Gaps 36;

QY 1 TCAGAAATAGTGTGATGAGTTAAATAACAATGTGGAAGCTGTTCCCGAGACCTCTCTT 60
DB 121355 TCAGAGTAGTATTGCTGAGTTAATACACATACAGTCGTGTTCCGATACCTCTCAG 121296
QY 61 CTTCTTTTCAGTACATGATTTGCTGGAAGACCTGTGGCAAAAGAGGCAAGCAAGTAA 120
DB 121295 TCACCTTTTCAGTACCTGACTGCTTAAAGGACATGTGGAAGCAAGCAAGTCAAGTTA 121236
QY 121 RAGTAGATATCCT-----TGTCCTTCCATTCGATTTTCAGCTATAAATG 168
DB 121235 AAGTAGGCTTCTCAGGGGGGGGGGGGGTCTACCTAACCTCTTTTATAAATG 121176
QY 169 GAACGTTTAGACTGCACGAGAAATGCATGGTTGTGAGAAATTA----- 212
DB 121175 AAACGTGTTAACTGTGTATACACAGTCTCTCCCATTTAAACCATCAGGGTTAACGACAC 121116
QY 213 -----ACATTTCTGGGTTAGTGAATAGCAATCATAGCTTTTGGGCACCTTCCCTG 264
DB 121115 TTGCTCTGATTTCTGTCAGTCCCTGCTCTCTCTGCTGCTCCCTCAGGCACACTCATG 121056
QY 265 GAACCTGCCAGATAAGCACTATTGAGCTCTTATTCCTCAGTCTGACATCAGCAAGTGTGAT 324
DB 121055 CAGTTTGGCCAGACAAG-ACCTCCTCAAGTGTGTAGCCCTTGTGCGCACTTTCACACTACAATGAC 120997
QY 325 TTCTCTATGAAAAATTTCTACTATGACTCTCTTATTTTAAAGTATACAAGAAACTTGTGACTCA 384
DB 120996 TTCTCTTGAGAAATGCTCACTCGGGCTTCTTACTGTGAGTCTATTAACAATTTGTGACTCA 120937
QY 385 GAAGATAATATTTACAGAGTGGAAAAAACCCCTAGCATTTATAGTTTAAACATTGAGG 444
DB 120936 GAATATAATATGCATAG-----GATAAAAACTTAGTATTTGTAAGGTACTCAAT----- 120888
QY 445 TTTTGAATGAGAGAGCTTATCCATAATATATTCAAATGTGTGTGAGTATGACACCTAAC 504
DB 120887 -----ACCTGAGAGTACTGAATATGAGAACCATCAGTGT-----GATATATCTGTAG 120838
QY 505 CTGTGATCTTGGAGTCAAGAAATGTTGAGTGTGCTGTTGAGTGTGCTGAGGAACAGCTAG 564
DB 120837 CAGTGACATCTGGCATACAAATCTTGAGGGCTACGGGCTGTTGGTGGGAATCAAGTTGG 120778
QY 565 TCGGTGAGCTGGCACAGGCATCTCAGTGAGTAGCATACCCACAGTTGGAAATTTTCAA 624
DB 120777 GGTGTCA-CATGGCACAGGCTTTTCAGAGAGCAGCACACCAAGGGTAGACACAGTTA 120719
QY 625 AGAATCAAAGGATCATGACA--TCTTATTAATTTCAAGGTTTCTGCTATACTTATGTA 682
DB 120718 GGCATCAAAATGAATTCAGAAGTGTCTTTGTAATCCTTAAGGCTTTGCAAGTACCCATGGG 120659
QY 683 AATGATAAATAAATCAAGCATATCCACTCTGTAAGATTGAACTTCTCAGATGGAAGACC 742
```


2851 -----CAGTTAAGAGGAGTTTCTCTGACAAATAATGTGAGCT 2887
Db 118883 ACATCCAGTACACACATATGCACACTAAATAAATGTGTTAAATGTGCGAGCT 118824
Qy 2888 AAGATAAGGGAAGAAAGCAGTTTTAGTTTCCCAAAATTTATTTTGGTGAGAGATT 2947
Db 118823 AAGGATCTGGAGGAA--GCTGTGCCCGTTGCCCTTCTTACTTTTGCTGATGGCTT 118766
Qy 2948 TTATTTTGTCTTTTATTTAGTGGAATATTTGGCTTGGAATGATAAATTTAGCCAGCTTAC 3007
Db 118765 ATA---TCCATTCTCTGTAGTGGAATATTTGGCTTGCAATGATAAGATTAGCCAGCTCAC 118709
Qy 3008 CAGGATGGACCCACAGAACTTTTCATAGAAATGGAGGACTGGAAGGAGACAAAGTAAA 3067
Db 118708 CAGGATGGACCCACAGAACTTCTCATTTGAAATGGAGGACTGGAAGGAGACAAAGTAAA 118649
Qy 3068 GGCTCACTATGAGGATTTCATCTGTACAGAAATGAAGCCAAATAATACCAGATCTCAGTGAA 3127
Db 118648 GGCACATTATGGAGCTTCACGGTACAGAACGAGCCAGCAATACCAAGTCTCAGTGAA 118589
Qy 3128 CAAATACAGAGGAACAGCGGTAATGCTCTCATGTGAGATGAGCATCTCAGCTGAGGAGA 3187
Db 118588 CAAATACAGAGGAGCGGTGGCAACGCCCTCATGGACGGGGCTCTCAACTGGTGGGGA 118529
Qy 3188 AACAGGACCATGACCATTCACAAACGGCATGTTCTTCAGCAGCTATGACAGAGCAATGA 3247
Db 118528 GAAACAGACCATGACCATTCACAAACGGCATGTTCTTCAGCAGCTATGACAGGAGCAACGA 118469
Qy 3248 CGGCTGGTATGTGGCACTCTTTGCTCTCG--CTTTAAATAATCACATAATCATTTACT 3306
Db 118468 CGGCTGGTAAAGCACCAAGGCTTTAGCTCTGCTGCTTGAATGATGCCACGAATCGGATTGTC 118409
Qy 3307 CAGAAATCAATA-----CAATATTTTAAATAGCTACCACTCTCTGGGCACTTA 3354
Db 118408 TAGCACTGTGGGGGTACCGAAATGTTCTTAAACAACTGCCAGACACTTGGACACTTCAAGTGT 118349
Qy 3355 --CTGTACGCCACTGCTTAACTCTTTATGATCACTCGAAGGATTTCAACTATAAGG 3412
Db 118348 ATCAGGAGCCCCATCCCGAACTCTTACGTATCAACGGGAGCATCCAGCTATAAGA 118289
Qy 3413 TAGACATCTTTATTTCTATTACAGATGAGATTTAGAGAGATTACGTGATTTGTCCAAT 3472
Db 118288 CAGACATGGTTAGTTTCACTTTTCAGATGAAGTTTAGAGATACGATTTATCCC----- 118236
Qy 3473 GTCAACAACTACCCAGAGATAAACTAGAAATTTAGACAGATTTCTTCTGAAATAAGA 3532
Db 118235 -TAAGACAGCTACCCAAAGATAGACCTGTATCTATAGCAACAAC-----TCTGAGTTTGT 118182
Qy 3533 GCATTTAGATAAATACCTATATCTCTATATTTCTAAAGTGTGTGAAAATTTTCATTTTC 3592
Db 118181 GTCTCTATATACATCTGTGTTTTCATTCATCAAGCATACGGAAGCCATCATTTCA 118122
Qy 3593 ATTTCCAGGGTTCTCTGATACATAAGGGTTGTAAAAGCTATTATTCAGATATAAGTAAACA 3652
Db 118121 TATCTCAGGTCT-----CTCATACTACAGTATCAGTAAAG 118085
Qy 3653 AACACAGTCCCTAGATGGATTG--GCACAAAGCCCGAGTTATCTCTCTTCTGCTATAGG 3711
Db 118084 AATTTCAGTCCCTTCATGATCAGATAGACCCAGCTGTTTCCCTTTTCATGCTATGGG 118025
Qy 3712 GCACAGGAGTCTTTGGTGTATTTAGTGTGACTGTATGATAGCACCCCAAGGAAGACTA 3771
Db 118024 GCAGAGAAGAACTTGCTGTATTTG---GGCTTTATACATCAGATTTCAGGGGAAGAGGT 117968
Qy 3772 CTGTGCACAGAGTGTAGCAGTCTTTTATGGGTAATCTGCAAAACCGCTTAACCTGACCACCG 3831
Db 117967 CCACACACAGTGGAGTACATACTTTTCATGCGGGTACTCTCCACGGAATGATTCATGG 117908
Qy 3832 TAGTCTCTGTTCTAATAACGCCAAACA--CATTTTCTTCAGGTAAACATCATGCCAG 3889
Db 117907 TGGCACTTGTGGATCATAACTCACACTCATCTCTCTTTTCAGGGTAACTACCGATCCAAAG 117848

Qy 3890 AAAACAGTGTCTTAAAGAGACGGTGGTGGATGTGGTATATAGATGTTCATGCAGCCAA 3949
Db 117847 AAAACAGTGTCTTAAAGAGATGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 117788
Qy 3950 TCCAAACGGCAGATACTACTGGGTGGACAGTACACCTGGGACATGGCAAAAGCATGGCAC 4009
Db 117787 TCCAAATGGCAGATACTACTGGGTGGCTTTTACAGCTGGGACATGTCCAAAGCAGGCAC 117728
Qy 4010 AGATGATGGTGTAGTATCGATGAATGGAGGGTTCATGTTCTCAATGAGGAAGATGAG 4069
Db 117727 AGATGATGGTGTGGTGTGGATGAATGGAGGGATCGTGGTACTCAATGAGGAGATGTC 117668
Qy 4070 TATGAAGATCAGGCCCTTCTTCCACAGCAATAGTCCCAATACCTAGATTTTGTCTCT 4129
Db 117667 TATGAAGATCAGGCCCTTCTTCCACAGCAATAGTCCCAATACCTAGATTTTGTCTCT 117608
Qy 4130 CTGTATGTGACAACTTTTGTACATTATGTTTATTTGGAATTTTCTTTCATACATTATTT 4189
Db 117607 TTCTATGT-CTGACATTTTTCATATTATGTTTATTTGGAATTTTCTTGTACATTATA-T 117550
Qy 4190 CCTCTAAACTCTCAAGCAGACGCTGAGTGTGACTTTTGGAAAAAGTATAGGATAAATA 4249
Db 117549 CCACATAAGCTCAAAACAATGTTGAGTGTGCTCTTTTGGAAACAATATAGGATAAAGAC 117490
Qy 4250 CATTAAATAGCACATGATTTTCTTTTCTTTCTTCTTCTTCTTCTTCTTCTTCTTCT 4292
Db 117489 ATTGAACAGCACCTAGATTTCTTTTGTATTTCTTCTTCTTCTTCTTCTTCTTCT 117447

RESULT 13
HUMFBRB3
LOCUS HUMFBRB3 763 bp DNA linear PRI 24-APR-1996
DEFINITION Human fibrinogen beta-chain gene, 3' end.
ACCESSION J00132
VERSION J00132.1 GI:182433
KEYWORDS beta-fibrinogen; fibrin; fibrinogen; glycoprotein.
SEGMENT 3 of 3
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 763)
AUTHORS Chung,D.W., Que,B.G., Rixon,M.W., Mace,M. Jr. and Davie,E.W.
TITLE Characterization of complementary deoxyribonucleic acid and genomic deoxyribonucleic acid for the beta chain of human fibrinogen
JOURNAL Biochemistry 22 (13), 3244-3250 (1983)
MEDLINE 83283433
PUBMED 6688356
COMMENT Original source text: Homo sapiens (clone: lambda-HI-beta-18K.)
(tissue library: T.Maniatitis) DNA.
FEATURES
source location/Qualifiers
1..763
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/map="4q28"
/clone="lambda-HI-beta-18K."
/tissue_lib="T.Maniatitis"
Join(J00130.1:17..144,J00131.1:1..222,1..466)
/gene="FGB"
<1..466
/gene="FGB"
/note="G00-119-130"
/number=8
<1..463
/gene="FGB"
/note="G00-119-130"
/number=8
<1..202
/gene="FGB"
/note="G00-119-130"
/number=8
<1..199

```
/gene="FGB"
/note="G00-119-130"
/number=8
<1..133
/gene="FGB"
/note="G00-119-130"
/number=8
<1..33
/gene="FGB"
/note="precursor"
/codon_start=1
/product="beta-fibrinogen"
/protein_id="AAA98116.1"
/db_xref="GI:182436"
/db_xref="GDB:G00-119-130"
/translation="MKIRPPFPQ"
<1..30
/gene="FGB"
/product="beta-fibrinogen"
/note="G00-119-130"
ORIGIN About 6 kb after segment 2; chromosome 4q31.
Query Match 14.0%; Score 683; DB 9; Length 763;
Best Local Similarity 98.7%; Pred. No. 1.9e-130;
Matches 741; Conservative 0; Mismatches 5; Indels 5; Gaps 5;
QY 4071 ATGAAGATCAGCCCTCTTCCACAGCAATAGTCCCAATACGAGATTTTGTCTCTC 4130
Db 1 ATGAAGATCAGCCCTCTTCCACAGCAATAGTCCCAATACGAGATTTTGTCTCTC 60
QY 4131 TGTATGTGACAACTTTTGTACATATGTTATTTGGAATTTTCTTTTACATATATTC 4190
Db 61 TGTATGTGACAACTTTTGTACATATGTTATTTGGAATTTTCTTTTACATATATTC 120
QY 4191 CTCTAAACTCTCAGACGAGTGAGTGAGCTTTTGAAAAGCTATAGGATAATATAC 4250
Db 121 CTCTAAACACACAGACGAGTGAGTGAGCTTTTGAAAAGCTATAGGATAATATAC 180
QY 4251 ATTTAAATAGCAGATGATTTCTTTTGTGTTTCTTCAATTTCTCTCTCACCAAGAGTA 4310
Db 181 ATTTAAATAGCAGATGATTTCTTTTGTGTTTCTTCAATTTCTCTCTCACCAAGAGTA 239
QY 4311 ACAAAAGTATAGTTTGTGACAGAGTGGTGTTCATATTTTCAATTTTCAATTTGCGAGA 4370
Db 240 ACAAAAGTATAGTTTGTGACAGAGTGGTGTTCATATTTTCAATTTTCAATTTGCGAGA 299
QY 4371 ATTTTCAATAGGAAGAGGGTCTTTTATCTCTGTCGTAGGAAACCATGACGGAAGG 4430
Db 300 ATTTTCAATAGGAAGAGGGTCTTTTATCTCTGTCGTAGGAAACCATGACGGAAGG 358
QY 4431 AAAAAGTATGTTTAAAGTCCACCTTTTAAACTATATTTATTTATGAGGATCTGTCAA 4490
Db 359 AAAAAGTATGTTTAAAGTCCACCTTTTAAACTATATTTATTTATGAGGATCTGTCAA 418
QY 4491 AGAAAGTCTCAAAAGATTTTATTAATTAACACGACTCTGTTGCAATAGTAAATGTTT 4550
Db 419 AGAAAGTCTCAAAAGATTTTATTAATTAACACGACTCTGTTGCAATAGTAAATGTTT 478
QY 4551 TCTTCTTTTGTAAATCCACATTCATGAGTTAGGCTTTGCACTTTGTAAGGAAGGAGAAG 4610
Db 479 TCTTCTTTTGTAAATCCACATTCATGAGTTAGGCTTTGCACTTTGTAAGGAAGGAGAAG 537
QY 4611 CGTTCACAACTCAATAGCTAATAAACCGGTCCTTGAATATTTGAAGATTTAAATCTGA 4670
Db 538 CGTTCACAACTCAATAGCTAATAAACCGGTCCTTGAATATTTGAAGATTTAAATCTGA 596
QY 4671 CTCTAGGACGGCAGCGGTGCTCAGACTATTAATCCCAACACTTTGGAGGCTGAGGCGG 4730
Db 597 CTCTAGGACGGCAGCGGTGCTCAGACTATTAATCCCAACACTTTGGAGGCTGAGGCGG 656
QY 4731 CGGGTCAAGAGTCAAGAGCTCAAGACAGGCTGACCAATATGTTGAAACCCCATCTCTA 4790
Db 657 CGGGTCAAGAGTCAAGAGCTCAAGACAGGCTGACCAATATGTTGAAACCCCATCTCTA 715
```

```
QY 4791 CTAAAAATACAAAAATTAGCCAGCGTGGT 4821
Db 716 CTAAAAATACAAAAATTAGCCAGCGTGGT 746

RESULT 14
BC036007 1792 bp mRNA PRI 20-SEP-2002
LOCUS Homo sapiens, clone IMAGE:4767275, mRNA.
DEFINITION BC036007
ACCESSION BC036007
VERSION BC036007.1 GI:23241682
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1792)
Direct Submission
Strausberg,R.
TITLE
JOURNAL
REMARK
COMMENT NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: gcgaps@mail.nih.gov
Tissue Procurement: CLONTECH
CDNA Library Preparation: CLONTECH Laboratories, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome
Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: http://www.shgc.stanford.edu
Contact: (Dickson, Mark) mcd@paxil.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
R. M.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IPAL Plate: 41 Row: o Column: 16.

FEATURES
source
1..1792
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4767275"
/tissue_type="Liver"
/clone_lib="NIH MGC_76"
/lab_host="DH10B"
/note="Vector: pDNR-LIB"

ORIGIN
Query Match 13.4%; Score 655.4; DB 9; Length 1792;
Best Local Similarity 99.0%; Pred. No. 8.5e-125;
Matches 670; Conservative 0; Mismatches 6; Indels 1; Gaps 1;
QY 3869 CAGGTTACATCAGATCCAGAAAACAGTTCCTAAAGAACGCGTGGTGGTGGTGA 3928
Db 1072 CTGGTTAATCATCAGATCCAGAAAACAGTTCCTAAAGAACGCGTGGTGGTGGTGA 1131
QY 3929 TAATAGATGTCATGCGCCAAATCCAAACGGCAGATACCTACTCTGGGTGGACAGTACACCTG 3988
Db 1132 TAATAGATGTCATGCGCCAAATCCAAACGGCAGATACCTACTCTGGGTGGACAGTACACCTG 1191
QY 3989 GGACATGCAAGGATGCGACAGATGATGGTGTAGTATGATGATTAATGGAAGGGGTCTATG 4048
Db 1192 GGACATGCAAGGATGCGACAGATGATGGTGTAGTATGATGATTAATGGAAGGGGTCTATG 1251
QY 4049 GTACTCAATGAGGAGATGAGTATGAAGATCAGGCCCTTCTTCCACAGCAATAGTCCCC 4108
Db 1252 GTACTCAATGAGGAGATGAGTATGAAGATCAGGCCCTTCTTCCACAGCAATAGTCCCC 1311
QY 4109 AATAGTAGATTTTGTCTCTCTGTATGTGACAACTTTTGTATATTATTATGGA 4168
```

```

|||||
1312 ANACGTAGATTTTGGCTCTCTCTGTATGTGACAAATTTTGTACATATATGTATGGAA 1371
QY TTTTCTTTTCATACATATATTTCTCTTAAACTCTCAAGACAGTGTGAGTGACATTTTGG 4228
Db TTTTCTTTTCATACATATATTTCTCTTAAACTCTCAAGACAGTGTGAGTGACATTTTGG 1431
QY AAAAAAGTAGATAGATAAATTCATATAAATAGACATGATTTTCTTTTCTTTCTTCATT 4288
Db AAAAAAGTAGATAGATAAATTCATATAAATAGACATGATTTTCTTTTCTTTCTTCATT 1491
QY TCTCTTGTCTACCCAGAAAGTAACAAAAGTATAGTTTTCACAGAGTGTGTTTCATAATT 4348
Db TCTCTTGTCTCA - CCAAGAAAGTAACAAAAGTATAGTTTTCACAGAGTGTGTTTCATAATT 1550
QY TCAGTCTCAGTGTGATGTCGAGAAATTTCAAATAAGAAAGAGGGTCTTTTATCCTTGTG 4408
Db TCAGTCTCAGTGTGATGTCGAGAAATTTCAAATAAGAAAGAGGGTCTTTTATCCTTGTG 1610
QY TAGGAAAACCATGACGGAAGGAAAGAACTGATGTTTAAAGTCCACTTTTAAACCTATAT 4468
Db TAGGAAAACCATGACGGAAGGAAAGAACTGATGTTTAAAGTCCACTTTTAAACCTATAT 1670
QY TTATTATAGTAGATCTGTCAAGAAAGAACTTCCAAAAGAGATTATTAATTAACACGACT 4528
Db TTATTATAGTAGATCTGTCAAGAAAGAACTTCCAAAAGAGATTATTAATTAACACGACT 1730
QY CTGTTGCAATAGTTAA 4545
Db CTGTTGCAATAAAAAA 1747

RESULT 15
LOCUS HUMFBRB Human fibrinogen beta-chain mRNA, partial cds. PRI 08-NOV-1994
DEFINITION Human fibrinogen beta-chain mRNA, partial cds.
ACCESSION J00129
VERSION J00129.1 GI:182429
KEYWORDS beta-fibrinogen; fibrin; fibrinogen; glycoprotein.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1883)
Chung, D.W., Que, B.G., Rixon, M.W., Mace, M. Jr. and Davies, E.W.
Characterization of complementary deoxyribonucleic acid and genomic
deoxyribonucleic acid for the beta chain of human fibrinogen
Biochemistry 22 (13), 3244-3250 (1983)
83283433
6680356
Original source text: Human liver, cDNA to mRNA, clone
pHI-beta-[1-6].
The authors identified three potential translation initiation
codons. Two of these codons were located upstream of position 1,
while the third is located at positions 19-21. The exact
initiation codon was not confirmed.
Location/Qualifiers
1..1883
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/map="4q28"
1..1883
/gene="FGB"
<1..1883
/gene="FGB"
/note="b-fibrinogen mRNA (alt.) [2]"
<1..1621
/gene="FGB"
/note="b-fibrinogen mRNA (alt.)"
<1..1618
/gene="FGB"
/note="b-fibrinogen mRNA (alt.)"

```

```

mRNA <1..1552
/gene="FGB"
/note="b-fibrinogen mRNA (alt.)"
CDS <1..1452
/gene="FGB"
/note="beta-fibrinogen precursor"
/codon_start=1
/protein_id="AAAS2429.1"
/db_xref="GI:182430"
/translation="FHLKTKMKHLHLLLVKSKQGVNDNEEFFSARGHRPLDKK
REAPLGPAPPISGGYRARPAPAKAAATQKVKERKAPDAGGLHADPLGLVLCPTGC
LOEALLQPERPINSVDLNNVEAVSOTSSSFQYVLLKDLWKQKQKOVKQDNEV
VNEYSSELEKHQYIDETVNSNIATNLRLSILNLSKIQLESDDYSAOMEYCEPT
CTVGSNIPIVSGKEBEIIRKGETSEMILIQDSSVVPYRVCMDMNTENGWTVIQN
RQDGNIDFVRKWDPIYKQGVNATNTDKNYCGLPGEYWLGNDKIQISQTRMGPTBELL
EMEDGKDKVAHYGGFTVQNEANKYQISVKNYRGTAGNALMDGASOLMGENRTWTH
NGMFSTYDRONDGLTSDPRKQSKEDGGWYNRCHAAANPNRYVYMGQYTDWMAK
HGTDDGVVMNWKSGWSMRKXMKIRPFPPQ"
sig_peptide <1..66
/gene="FGB"
/note="beta-fibrinogen signal peptide"
mat_peptide 67..1449
/gene="FGB"
/product="beta-fibrinogen"
ORIGIN 114 bp upstream of TaqI site; chromosome 4q31.
Query Match 13.4%; Score 653.4; DB 9; Length 1883;
Best Local Similarity 99.7%; Pred. No. 2.2e-124;
Matches 665; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
QY 3869 CAGGTTAACTACATCCAGAACAGTGTCTTAAAGAGACGGTGTGGTGGTGA 3928
Db TTTTCTTTTCATACATATATTTCTCTTAAACTCTCAAGACAGTGTGAGTGACATTTTGG 4228
QY 3929 TAATAGATGTCTATGACAGCAATCCAAAAGGAGATCTACTGGGTGGACAGTACACCTG 3988
Db TAATAGATGTCTATGACAGCAATCCAAAAGGAGATCTACTGGGTGGACAGTACACCTG 1337
QY 3989 GGACATGGCAAGCATGGCACAGATGATGGTGTAGTATGATGAATTTGGAAGGGTCAATG 4048
Db GGACATGGCAAGCATGGCACAGATGATGGTGTAGTATGATGAATTTGGAAGGGTCAATG 1397
QY 4049 GTACTCAATGAGGAAGATGAGTATCAAGATCAGGCCCTCTTCCACAGCAATAGTCCCC 4108
Db GTACTCAATGAGGAAGATGAGTATCAAGATCAGGCCCTCTTCCACAGCAATAGTCCCC 1457
QY 4109 AATAGCTAGATTTTCTCTCTGTATGTGACAAATTTTGTACATATATTTTGGAA 4168
Db AATAGCTAGATTTTCTCTCTGTATGTGACAAATTTTGTACATATATTTTGGAA 1517
QY 4169 TTTTCTTTTCATACATATATTTCTCTTAAACTCTCAAGACAGTGTGAGTGACATTTTGG 4228
Db TTTTCTTTTCATACATATATTTCTCTTAAACTCTCAAGACAGTGTGAGTGACATTTTGG 1577
QY 4229 AAAAAAGTAGATAAATTCATATAAATAGACATGATTTTCTTTTCTTTCTTCATT 4288
Db AAAAAAGTAGATAAATTCATATAAATAGACATGATTTTCTTTTCTTTCTTCATT 1637
QY 4289 TCTCTTGTCTACCCAGAAAGTAACAAAAGTATAGTTTTCACAGAGTGTGTTTTCATAATT 4348
Db TCTCTTGTCTCA - CCAAGAAAGTAACAAAAGTATAGTTTTCACAGAGTGTGTTTTCATAATT 1696
QY 4349 TCAGTCTCAGTGTGATGTCGAGAAATTTCAAATAAGAAAGAGGGTCTTTTATCCTTGTG 4408
Db TCAGTCTCAGTGTGATGTCGAGAAATTTCAAATAAGAAAGAGGGTCTTTTATCCTTGTG 1756
QY 4409 TAGGAAAACCATGACGGAAGGAAAGAACTGATGTTTAAAGTCCACTTTTAAACCTATAT 4468
Db TAGGAAAACCATGACGGAAGGAAAGAACTGATGTTTAAAGTCCACTTTTAAACCTATAT 1816
QY 4469 TTATTATAGTAGATCTGTCAAGAAAGAACTTCCAAAAGAGATTATTAATTAACACGACT 4528

```

Db 1817 TTATTTATGTAGATCTGTCAAAGAAAACCTTCCAAAAAGATTATTATTAATAACGAGACT 1876
Qy 4529 CTGTTGC 4535
| | | | |
Db 1877 CTGTTGC 1883

Search completed: November 23, 2004, 04:31:25
Job time : 20238.7 secs

Result No.	Score	Query		DB	ID	Description
		Match	Length			
C	1	260	12.9	1056	4	BM470567
	2	256.8	12.8	497	1	AL589495
	3	253.8	12.6	834	6	CD357520
	4	253.2	12.6	565	6	CA426144
	5	253	12.6	576	8	AQ24282
	6	252	12.5	704	5	BX645059
	7	250.2	12.4	635	8	B91104
	8	248.4	12.4	706	5	BUE18120
	9	247.6	12.3	680	7	CF887062
	10	247	12.3	699	5	BUE77785
C	11	247	12.3	753	6	CF890780
	12	246.8	12.3	775	6	CF898598
	13	246.6	12.3	838	3	BC039515
	14	246.6	12.3	1060	4	BM469250
	15	246.4	12.3	733	6	CB963738
	16	245.2	12.2	543	8	AQ514325
	17	245	12.2	895	5	BQ437946
	18	245	12.2	602	2	AW965008
	19	244.8	12.2	635	1	AL691639
	20	242	12.0	713	9	AG011265
C	21	241.6	12.0	673	9	AG175165
	22	241.6	12.0	3143	3	HSM805999
	23	241.4	12.0	642	5	BN959857
	24	241.2	12.0	537	2	BF875683

```

Db      526  GGTAGGCTAGTTGGTGTGATTATACACTGTAATCCACAGCACTTTGGGAGGCGCTAGGTGGG 467
Qy      348  TAGATCACTTGGAGCGCGGAGTTTGAGACCAAGCCTGSCCAACATGCTGTAACCCCATCTG 407
Db      466  TGGATCGCTTGGAGTCAAGGAGTTTGGAGACCAAGCCTGSCCAACATGCTGTAACCCCATCTG 407
Qy      408  TACTAAAAATACAAAAGTTAGTCTGGCGTGTGTGTAGATGCTCTAGTCCAGCACTACTTG 467
Db      406  TACCAAAAATATAAAAATTTAGCCAGGTGCAGTGCAATGTGCCCGTAGTCCAGCACTACTTG 347
Qy      468  GGAGGCTGAGGCATGAGAATCGCTTGGAGCCAGCCTGGGCAATACAGCAAGACCCCGTCT 527
Db      346  GGAGGCTGAGGTGGGAGATCACTTGTATCCAGCCTGCTCAACGCTGGCGAAACCCACCT 287
Qy      528  CTACAAATATAAATAAATAGTTGGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 587
Db      286  CTAC-----CAAAAATACAAAAATTTAGCTGGCATGTTGTGTGTGTGTGTGTGTGTGTGTGTGT 231
Qy      588  CTAGGAGGCTGAGATGAAGGATTTGTTGAGCCTTGGGAGGTCAAGGCTGCAGTGAGCGG 647
Db      230  CTTAGGAGGCTGAGCATGAGATCACTTTGAACCCAGGAGACAGAGGTCAAGTGAGCGG 171
Qy      648  AGATGGGCCACTGCCTCCAGCCTGGGCAACAGAGTGAGACCCCTGTCTCAGAAAAAAA 707
Db      170  ATATTGTGCCACTGCCTCCAGCCTGGGCAACAGAGTGAGACTCTGTCTCAAAAAAATTT 111
Qy      708  AAAAAAAGGAGGAGAGAGAG 733
Db      110  AAATAAATAATTAATTAATGAAG 85

```

```

RESULT 2
AL589495/c
LOCUS      DKEP451G0619 r1 451 (synonym: hicc1) spinal cord Homo sapiens cDNA
DEFINITION clone DKEP451G0619 5', mRNA sequence.
ACCESSION AL589495
VERSION    AL589495.1 GI:13243267
KEYWORDS   EST.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

```

```

REFERENCE 1 (bases 1 to 497)
AUTHORS   Bloeker,H., Boeche,M., Brandt,P., Mewes,W., Weil,B. and
            Wiemann,S.
TITLE      EST (Bloeker,H., Boeche,M., Brandt,P., Mewes,H.W., Weil,B. and
            Wiemann,S.)
JOURNAL    Unpublished (1999)
COMMENT    Contact: MIPS
            MIPS
            Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
            This is the 5' sequence of the clone insert
            Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
            Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
            sequenced by GBF (National Research Centre for Biotechnology Ltd.,
            Braunschweig/Germany) within the cDNA sequencing consortium of the
            German Genome Project.
            No si sequence available.
            This clone (DKFZp451G0619) is available at the RZPD in Berlin.
            Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
            Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.

```

```

FEATURES
source
1..497
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKFZp451G0619"
/tissue_type="human spinal cord"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="451 (synonym: hicc1) spinal cord"
/notes="Vector: pSport1; Site_1: NotI; Site_2: SalI"

```

ORIGIN

```

Query Match      12.8%; Score 256.8; DB 1; Length 497;
Best Local Similarity 78.3%; Pred. No. 8.5e-32;
Matches 351; Conservative 0; Mismatches 82; Indels 15; Gaps 3;

Qy      292  GGCGGGGTGTGGTGTCTATCCCTCTAATCCAGCACTTTGGGAGGCGCAAGTGTGAGTAGA 351
Db      487  GGCGGGCGGGTGGTGTCTATCCCTCTAATCCAGCACTTTGGGAGGCGCGAGTGTGGTGGG 428
Qy      352  TCATTGAGGCGGGAGTTTGAGCACCAGCCTGGCCCAACATGTTGTAACCCCATCTGTACT 411
Db      427  TCACCTGAGGTCAAGAGCTCAAGACCAGCCTGGTTAACTATGGCCAAACCCCGTCTCTACT 368
Qy      412  AAAAATACAAAGTTA-GCTGGGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 470
Db      367  AAAAATACAAATTTAGCCGGGCGGTGGCTCAAGCCTGTATCCAGCACATTTGGGA 308
Qy      471  GGCTGAGGCATGGAATC-----GCTTGAGCCAGCCTGGGCAATACAGCAAG 518
Db      307  GGCCGAGGCGGGCGGATCACAAAGGTCAAGAGATCGAGACCATCTCTGGCTAACCGGTGAA 248
Qy      519  ACCCGCTCTCTACAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 578
Db      247  ACCCGCTCTCTACTA--AAAATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 190
Qy      579  TCCTAGCTGTGTAGGAGGCTGAGATGGAAGATTGCTTGAGCCTGGGAGGTCAAGGCTGC 638
Db      189  TCCAGCTGTCTCCGAGGCTGAGGCGAGAGATGCGCATGGAACCTGGGAGCGGAGCTTAC 130
Qy      639  AGTGAGCGAGATGGCGCCACTGCACCTCCAGCCTGGGCAACAGAGTGAGACCTGTCTCA 698
Db      129  AGTGAGCTGAGATCGCGCCACTGCACCTCCAGCCTGGGCGCAGAGCAAGACTCCGTTTCA 70
Qy      699  GAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 726
Db      69  AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 42

```

RESULT 3

```

CD357520
LOCUS      CD357520
DEFINITION AGNCOURT 14248776 NIH MGC 187 Homo sapiens cDNA clone
            IMAGE:30401007 5', mRNA sequence.
ACCESSION CD357520
VERSION    CD357520.1 GI:31128931
KEYWORDS   EST.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

```

```

REFERENCE 1 (bases 1 to 834)
AUTHORS   NIH-MGC http://mgs.nci.nih.gov/
TITLE      National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL    Unpublished (1999)
COMMENT    Contact: Daniela S. Gerhard, Ph.D.
            Office of Cancer Genomics
            National Cancer Institute / NIH
            Bldg. 31 Rm10A07 Bethesda, MD 20892
            Email: cgapbs@mail.nih.gov
            Tissue Procurement: Dr. Michael Brownstein and Dr. Miklos Palkovits
            cDNA Library Preparation: CLONTECH Laboratories, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Agencourt Bioscience Corporation
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: NDCM181 row: a column: 16
            High quality sequence stop: 381.

```

FEATURES

```

Source
1..834
/organism="Homo sapiens"
/mol_type="mRNA"

```


/db_xref="taxon:9606"
 /clone="IMAGE:30401007"
 /lab_host="DH10B (T1 phage-resistant)"
 /clone_lib="NIH_MGC_187"
 /note="Organ: Blood vessels - aorta, basilar and artery;
 Vector: pDNR-LIB; Site 1: SfiI (ggccattatggcc); Site 2:
 SfiI (ggcgctggcc); 5' and 3' adaptors were used in
 cloning as follows: 5' adaptor sequence:
 5'-CACGGCCATTATGGCC-3' and 3' adaptor sequence:
 5'-ATTCTAGAGCCGCGGCGCCGACATG-dT(30)BN-3' (where B = A,
 C, or G and N = A, C, G, or T). Average insert size 1.4 kb
 (range 0.5-4.0 kb). 14/15 colonies contained inserts by
 PCR. This library was enriched for full-length clones and
 was constructed by Clontech Laboratories (Palo Alto, CA).
 Note: this is a NIH_MGC Library."

ORIGIN

Query Match 12.6%; Score 253.8; DB 6; Length 834;
 Best Local Similarity 69.1%; Pred. No. 2.3e-31;
 Matches 367; Conservative 0; Mismatches 152; Indels 12; Gaps 1;
 QY 276 AGTGTGAGGGAGGTGGCGGGTGTGGTCTATGCTCTTAATCCAGCATTTGGGA 335
 Db |||||
 QY 3 AGAGTAAGATATCCCTGGCGGGCAGGTGGCTGCACCTGTAATCCCGGCATTTGGGA 62
 Db |||||
 QY 336 GGCCAAAGTGTAGTATCACTTTGAGCGCGGAGTTTGACACAGCTGCGCCACATGGTG 395
 Db |||||
 QY 63 GGCTGAGGCGGATGGATCGCTTGAGGTGAGGAGTTTGACACAGCTGACCAATGGTG 122
 Db |||||
 QY 396 AAACCCCATCTTACTATAAAATACAAAGTTAGTGGCGTGGTGTAGATCCCTGTAGT 455
 Db |||||
 QY 123 AAATCCCATCTTACTATAAAATACAAATTTGGCGGGCGTGGTGGCTCAGCCCTGTAAT 182
 Db |||||
 QY 456 CCCAGTACTTGGGAGGTGGAGCATGAGATC-----GCTTGACCCAGCT 503
 Db |||||
 QY 183 CCAGCACTTTGGGAGGTGGAGCGAGGATCATGAGTCAAGACATTTGAGACCAACCAT 242
 Db |||||
 QY 504 GGGCAATACAGCAAGACCCGCTCTTACAAATAAAATACAAATAATTTAGTGGATGGT 563
 Db |||||
 QY 243 GGCTAACATGTTGAACCCCGCTCTCTAATAGAAATACAAATAATTTAGTGGCGTGT 302
 Db |||||
 QY 564 GGTGATGCTGTAGTCTTGTAGTGGAGGCTGAGATGGAAGATTTGTTGAGCCTG 623
 Db |||||
 QY 303 GGCGGGCATCTGTAGTCCAGCTACTCGGAGGCTGATGGAGGAGATGGCGTGAACCTG 362
 QY 624 GGAGTCAAGGTGAGTGGAGGAGATGGCCACTGCACTCCAGCTGGCCGACACAGAG 683
 Db |||||
 QY 363 GGAGCGGAGCTTGCAGTGGCGGAGATTGCGCTGCTGCACTCCAGCTGGACACACAGAG 422
 Db |||||
 QY 684 TGAGACCTGTCTCAGAAAAAATAAAAAAAGAGAGGAGAGACTCAAGCAC 743
 Db |||||
 QY 423 CGTGAATCTCTCAAAAGAACCCCATATATAAAAAAACAATGTGCGCGGCGCTCG 482
 QY 744 GCCCTCAGAGACTGTGAGGCCCTGCAAGTGTCTGCAAGCATGTGGCCCC 794
 Db |||||
 QY 483 GCCCTTCAAGAAGCTTCTAGACCATTCGTTTGGCGGCGGCGCCCC 533
 Db |||||

RESULT 4

CA426144/c
 LOCUS
 DEFINITION UI-H-DF0-bej-f-14-0-UI-s1 NCI CGAP DF0 Homo sapiens cdna clone
 UI-H-DF0-bej-f-14-0-UI 3', mRNA sequence.
 CA426144
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM Homo sapiens (human)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 565)
 NCBI-CCGAP http://www.ncbi.nlm.nih.gov/ncicgap.
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

JOURNAL
COMMENT

Tumor Gene Index
 Unpublished (1997)
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: Dr. Jose Mercuende
 cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: Clone distribution information can be obtained
 from Dr. M. Bento Soares, bento-soares@uiowa.edu
 The following repetitive elements were found in this CDNA
 sequence: 11-299, >ALU (matched complement) 217-435, >ALU (matched
 complement) 237-527, >SVA#Other
 Seq primer: M13 FORWARD
 POLYA=yes.

FEATURES

Location/Qualifiers
 1..565
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="UI-H-DF0-bej-f-14-0-UI"
 /tissue_type="Subchondral Bone"
 /dev_stage="Adult"
 /lab_host="DH10B (Life Technologies)"
 /clone_lib="NCI CGAP DF0"
 /note="Organ: Bone; Vector: pT73-Pac (Pharmacia) with a
 modified polylinker; Site 1: EcoR I; Site 2: Not I;
 NCI CGAP DF0 is a cDNA library containing the following
 tissue(s): Subchondral Bone. The library was constructed
 according to Bonaldo, Lennon and Soares, Genome Research,
 6:791-806, 1996. First strand cDNA synthesis was primed
 with an oligo-dT primer containing a Not I site. Double
 stranded cDNA was ligated to an EcoR I adaptor, digested
 with Not I, and cloned directionally into pT73-Pac
 vector. The oligonucleotide used to prime the synthesis of
 first-strand cDNA contains a library tag sequence that is
 located between the Not I site and the (dT)18 tail. The
 sequence tag for this library is GTTAAGCGTC.
 TAG_TISSUE=Subchondral bone
 TAG_LIB=UI-H-DF0
 TAG_SEQ=GTTAAGCGTC"

ORIGIN

Query Match 12.6%; Score 253.2; DB 6; Length 565;
 Best Local Similarity 78.8%; Pred. No. 3.1e-31;
 Matches 345; Conservative 0; Mismatches 78; Indels 15; Gaps 3;
 QY 291 TGGCGGGTGTGGTGGCTCTAATCCAGCACTTTGGGAGGCCAAGGTGAGTAG 350
 Db |||||
 QY 436 TGGCGGGCATGGTGGCTTATGCTGTATCTTGGCACTTTGGGAGGCCAAGTGGCGAG 377
 Db |||||
 QY 351 ATCACTGAGCGCGGAGTTTGAGACCGCTGGCCAAACATGGTGAACCCCATCTGTAC 410
 Db |||||
 QY 376 ATCACTGAGGTGAGAGTTTGAGACCGCTGGCCAAACATGGTGAACCCCATCTGTAC 317
 Db |||||
 QY 411 TAAATAACAAAGTTA-GCTGGGGCTGTGGTAGATGCTGTAGTCCAGCTACTTTGGG 469
 Db |||||
 QY 316 TAAATAACAAATTTAGCGCGGGCGGTGGCTCACGCTGTATATCCAGCACTTTGGG 257
 Db |||||
 QY 470 AGGCTGAGGCATGAGAATC-----GTTTGGCCCGAGCTGGGCAATACAGCA 517
 Db |||||
 QY 256 AGGCAGAGCGCGCGGATCAGAGGTGAGAGATCAAGACCATCTTGGCTAACCGGTGA 197
 Db |||||
 QY 518 GACCCCGTCTTACAAATAAAATACAAAAATTAGTTGGATGTGGTGGTGCCTGTA 577
 Db |||||
 QY 196 AACCCGCTCTACTA--AAAATAAAAAATTAGCGGGCTAGTGGCGGTGCTGTA 139
 Db |||||
 QY 578 GTCTAGTGTAGGAGGCTGAGATGAAGATGCTTTGAGCCTGGGAGGTCAAGGCTG 637
 Db |||||
 QY 138 GTCCGAGCTACTCGGAGGCTGAGGAGGAGATGGCATGAACCCCGGAGGACAGCTTG 79
 Db |||||
 QY 638 CAGTGAGCGGAGATGGCGCCACTGCTCCAGCTGGGCAACAGAGTCAGACCTGTCTC 697
 Db |||||

```

Db      78  CAGGAGCGGAGATTGCACCACTGCATCCAGCCTGGGCGGACAGAGGAGACTCCGTCTC 19
Qy      698  AGAAAAAAGAAAAA 715
Db      18  AAAAAAAGAAAAA 1

RESULT 5
LOCUS   AQ424282
DEFINITION  CITBI-E1-2574H5.TR CITBI-E1 Homo sapiens genomic clone 2574H5,
genomic survey sequence.
ACCESSION  AQ424282
VERSION    AQ424282.1 GI:4497548
KEYWORDS   GSS.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1  (bases 1 to 576)
AUTHORS    Zhao,S., Adams,M.D., Nierman,W., Malek,J., Shizuya,H., Simon,M. and
Venter,J.C.
TITLE      Use of BAC End Sequences from CalTech Libraries for Sequence-Ready
Map Building
JOURNAL    Unpublished (1997)
COMMENT    Other GSSs: CITBI-E1-2574H5.TF
Contact: Shaying Zhao, William Nierman, Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850
Tel: 301 838 0200
Fax: 301 838 0208
Email: hbs@tigr.org
Clones are available from Research Genetics (info@resgen.com). BAC
end search page:
http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
Seq primer: M13 Reverse
Class: BAC ends.

FEATURES             source
   source
1. 576
   /organism="Homo sapiens"
   /mol_type="genomic DNA"
   /db_xref="taxon:9606"
   /clone="2574H5"
   /sex="male"
   /cell_type="sperm"
   /clone_lib="CITBI-E1"
   /note="Vector: pBelBAC11; Site_1: EcoRI; Site_2: EcoRI;
CalTech Human BAC Library D"

ORIGIN
Query Match      12.6%; Score 253; DB 8; Length 576;
Best Local Similarity 77.0%; Pred. No. 3.4e-31;
Matches 352; Conservative 0; Mismatches 90; Indels 15; Gaps 3;

Qy      287  AGCGTGGCGGGGTGTGGTCTATCCAGCACTTTGGAGGCGCAAGGTGA 346
Db      38  ACGTGGCGGGCACAGTGGCTCACACCTGTAATCCAGCACTTTGGAGGCGCAAGGTGG 97

Qy      347  GTAGATCACTTGAGCCGGGAGTTTGAGACACAGCTGGCGCAACATGGTGAACCCCATCT 406
Db      98  GCGGATCACTCGAGTCAAGGATTTGAGACAGCCTGGCCAAATGGGCAACCCCTGTCT 157

Qy      407  GPTACTAAATAACAAAG-TTAGCTGGGGTGGTGTAGTGCCTGTAGTCCAGTACT 465
Db      158  CAACATAAATAATGAAGAATTAGCCGGGGTGGTGGTGCCTGTATCCAGCACTT 217

Qy      466  TGGGAGGCTGAGGCATGAATC-----GTTTGAGCCAGCCTGGGCAATACA 513
Db      218  TGGGAGGCGGAGCGGGTGAATCACAAAGTCAGGAGATCGAGACCATCTCGTTAAACAG 277

Qy      514  GCAAGACCCGCTCTACAAATAAATAAATAAATTAGTTGGATGTGGTGTGTCATGCC 573

```

```

Db      278  GTGAAACCTGCTCTCTACTA--AAATACAAAAAATTAGCCGGGAGTGGTTGTGGCGCC 335
Qy      574  TGTAAGTCTAGCTCTAGGAGGCTGAGATGGAAGGATTGCTTGAAGCTGGAGGTCAAG 633
Db      336  TGTAAGTCCAGCTACTCGGAGGCTGAGGAGGAGATGGCGTGAACCCGGGAGGTGGAG 395

Qy      634  GCTCAGTGAGCCGAGATGGGCGGCTGCACTCAGCTCGGCAACAGAGTGAGACCCCTG 693
Db      396  CTTGCACTGAGCGGAGATAGCACCACTGCATCTCCAGGCTGGGCGGACAGTGAGACTTTG 455

Qy      694  TCTCAGAAAAAAGAAAAAAGAAAAAAGAGGAGGAGAGA 730
Db      456  TCTCAAAAAAAGAAAAAATTAATAAAAAATGTAAAAAGA 492

RESULT 6
LOCUS   BX645059/c
DEFINITION  DXFP781H2437_r1 781 (synonym: hlccc4) Homo sapiens cDNA clone
BX645059
ACCESSION  BX645059
VERSION    BX645059.1 GI:34479392
KEYWORDS   EST.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1  (bases 1 to 704)
AUTHORS    Poustka,A., Albert,R., Moosmayer,P., Schupp,I., Wellenreuther,R.,
Mewes,H.W., Weil,B., Amid,C., Osanger,A., Fobo,G., Han,M. and
Wiemann,S.
TITLE      EST (Poustka,A., Albert,R., Moosmayer,P., Schupp,I.,
Wellenreuther,R., et al.)
JOURNAL    Unpublished (2003)
COMMENT    Contact: MIPS
MIPS
Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
This is the 5' sequence of the clone insert
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by DKFZ (German Cancer Research Center,
Heidelberg/Germany) within the cDNA sequencing consortium of the
German Genome Project.
No sl sequence available.
This clone (DKFP781H2437) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.

FEATURES             source
   source
1. 704
   /organism="Homo sapiens"
   /mol_type="mRNA"
   /db_xref="taxon:9606"
   /clone="DXFP781H2437"
   /dev_stage="adult"
   /lab_host="DH10B"
   /clone_lib="781 (synonym: hlccc4)"
   /note="Vector: pSport1_Sfi; Site_1: SfiIA; Site_2: SfiIB;
cDNA-collection"

ORIGIN
Query Match      12.5%; Score 252; DB 5; Length 704;
Best Local Similarity 78.4%; Pred. No. 4.6e-31;
Matches 345; Conservative 0; Mismatches 80; Indels 15; Gaps 3;

Qy      292  GGCGGGGTGTGGTGGCTCATCTTAATCCAGCACTTTGGAGGCGCAAGGTGAGTGA 351
Db      440  GCCAGGCAATGGTGGCTCACACCTGTAATCCAGCACTTTGGAGGCGGAGGTGAGTGA 381

Qy      352  TCACTTGAGCGGGAGTTTGAGACCAAGCTGGCCCAACATGGTGAACCCCATCTGTACT 411
Db      380  TCACTTGAGGTGAGGATTCGAGACCAAGCTGGCCCAACATGGTGAACCCCATCTGTACT 321

Qy      412  AAAATA-CAAAAGTTAGCTGGCGGTGGTGTAGATGCTGTAGTCCCGAGCTACTTGGGA 470

```

```

Db      320 ATAAATAGAAAAAATGGCTGGCGTGGTGGCTCAATGCTGTAATCCAGCAGCTTTGGGA 261
QY      471 GCGTGGAGCATGAGATC-----GCTTGAGCCAGCGCTGGGCAATACAGCAAG 518
Db      260 GGCAGGAGGGGGGGGATACGAGGTGAGGATCGAGAGCATCTGGCTTAACACCGTGAA 201
QY      519 ACCCGCTCTCTACAAATAAAATAACAAAAATAGTTGGATGTGGTGCATGCCCTGTAG 578
Db      200 ACCCGCTCTCTACTAAAAATACAAAAAATAGCCGGCTGGTGGCGGCGCTGTAG 141
QY      579 TCTAGCTGCTAGGAGGCTGAGATCGAAGGAT--TGCTTGAGCCCTGGGAGGTCAAGGCT 636
Db      140 TCCAGCTACTCGGAGGCTGAGGACGAGAGATGGTGGGTGAACCCGGGAGGGAGCTT 81
QY      637 GCAGTGAGCGGAGATGGCGCCACTGCACTCCAGCTGGGCAACAGAGTGAGACCCCTGTCT 596
Db      80 GCAGTGAGCGGAGATGCGACCACTGCACTCCAGCTGGGCGGACAGAGCGAGATCCATCT 21
QY      697 CAGAAAAAATAAAAAA 716
Db      20 CAAAAAATAAAAAA 1

```

```

RESULT 7
B91104/c
LOCUS   B91104
DEFINITION CIT-HSP-2173120.TF CIT-HSP Homo sapiens genomic clone 2173120,
genomic survey sequence.
ACCESSION B91104
VERSION   B91104.1 GI:2973584
KEYWORDS
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens

```

```

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 635)
Adams,M.D., Rounsley,S.D., Zhao,S., Field,C.E., Bass,S., Linher,K.,
Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H.,
Simon,M. and Venter,J.C.
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: mdadams@tigr.org
Clones are available from Research Genetics (info@resgen.com). BAC
end search page:
http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html
Seq primer: M13-21;
Class: BAC ends.

```

```

FEATURES
         Location/Qualifiers
            1..635
             /organism="Homo sapiens"
             /mol_type="genomic DNA"
             /db_xref="GDB:7104270"
             /db_xref="taxon:9606"
             /clone="2173120"
             /sex="Male"
             /cell_type="Sperm"
             /clone_lib="CIT-HSP"
             /note="Vector: pBelBAC11; Site_1: HindIII; Site_2:
HindIII"

```

```

ORIGIN
Query Match      12.4%; Score 250.2; DB 8; Length 635;
Best Local Similarity 73.5%; Pred. No. 9.3e-31;
Matches 352; Conservative 0; Mismatches 113; Indels 14; Gaps 2;

```

```

QY      245 CTACAGTTGTGATGTGGTACCATGCCAAGTGTGAGGGAGGCTGGCCGGTGTGGT 304
Db      635 CTGTAAAGCATAGAGAAACATCCACTTTTAGATGAAAAAGTTGTGCGCAATGCAGT 576
QY      305 GGCTCATGCTCTAATCCAGCACATTTGGAGGCCAAGGTGAGTAGATCACTTGAGGCCG 364
Db      575 GGCTCAAGCTGTATCCAGCACATTTGGAAAGCCGAGGTGAGCGGATCACCTGAGGTCA 516
QY      365 GGAGTTTGAGACAGCGCTGGCCAAACATGGTGAAGAACCCCATCTGTACTAAAAATACAAAAG 424
Db      515 GGAGTTTGAAACAGCGCTGGCCAAACATGGCAAAACCCCATCTCTACTAAAAATACAAAAG 456
QY      425 TTAGCTGGCGGTGGTGTAGATGCTGTAGTCCAGCTACTTTGGGAGGCTGAGGCATGAG 484
Db      455 TTAGCCAGGCACGGTGGTGGTCCCTGCAATCCAGCACTTTGGAGGCTGAGGGGGCA 396
QY      485 AATCGCTTG-----AGCCAGCGCTGGGCAATACAGCAAGACCCCGCTCTCTACA 532
Db      395 GATCACAGGTGAGGAGCAAGACCATCTCTGGCTAACACAGTGAACCCCATCTCTACT 336
QY      533 AATAAAATACAAAAAATTAGTTGGATGTGGTGGTGCATGCCCTGTAGTCTTAGCTGTAG 592
Db      335 A--AAATACAAAAAATTAGCCGGTGTGGTGGCGGCACCTGTAGTCCACTCTCAG 278
QY      593 GAGGCTGAGATCGAAGGATGCTTGAGCTGGAGGTCAAGCTGAGCTGAGTGGCCGAGATG 652
Db      277 GAGGCTGAGGAGGAGGAATGGATGAACCCGGGAGCGGAGCTTCGAGTGCAGAT 218
QY      653 GCGCCACTGCACTCCAGCGCTGGGCAACAGAGTGAGACCCCTGCTCAGAAAAA 711
Db      217 GCGCCACTGCACTCCAGCGCTGGGCAACAGTGTGAGACTCCGCTCAAAAAA 159

```

```

RESULT 8
B918120/c
LOCUS   B918120
DEFINITION UI-H-DF0-beu-1-08-0-UI.s1 NCI_CGAP_DFO Homo sapiens cDNA clone
UI-H-DF0-beu-1-08-0-UI 3', mRNA sequence.
ACCESSION B918120
VERSION   B918120.1 GI:23284335
KEYWORDS
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens

```

```

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 706)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Clone distribution information can be obtained
from Dr. M. Bento Soares, bent-soares@uiowa.edu
The following repetitive elements were found in this cDNA
sequence: 11-299, >ALU (matched complement) 217-435, >ALU (matched
complement) 351-572, >ALU (matched complement) 574-633,
>MIR#SINE/MIR
Seq primer: M13 FORWARD
POLYA=Yes.

```

```

FEATURES
         Location/Qualifiers
            1..706
             /organism="Homo sapiens"
             /mol_type="mRNA"
             /db_xref="taxon:9606"
             /clone="UI-H-DF0-beu-1-08-0-UI"
             /tissue_type="Subchondral Bone"
             /dev_stage="Adult"

```

```

ORIGIN
Query Match      12.4%; Score 250.2; DB 8; Length 635;
Best Local Similarity 73.5%; Pred. No. 9.3e-31;
Matches 352; Conservative 0; Mismatches 113; Indels 14; Gaps 2;

```

/lab host="DH10B (Life Technologies)"
 /clone lib="NCI CGAP DFO"
 /note="Organ: Bone; Vector: pT73-Pac (Pharmacia) with a
 modified polylinker; Site 1: EcoR I; Site 2: Not I;
 NCI CGAP DFO is a cDNA library containing the following
 tissue(s): Subchondral Bone. The library was constructed
 according to Bonaldo, Lennon and Soares, Genome Research,
 6:791-806, 1996. First strand cDNA synthesis was primed
 with an oligo-dT primer containing a Not I site. Double
 stranded cDNA was ligated to an EcoR I adaptor, digested
 with Not I, and cloned directionally into pT73-Pac
 vector. The oligonucleotide used to prime the synthesis of
 first-strand cDNA contains a library tag sequence that is
 located between the Not I site and the (dT)18 tail. The
 sequence tag for this library is GTTAAGCGTC.
 TAG TISSUE=Subchondral bone
 TAG_LIB=UI-H-DFO
 TAG_SEQ=GTTAAGCGTC"

ORIGIN

Query Match 12.4%; Score 248.4; DB 5; Length 706;
 Best Local Similarity 78.1%; Pred. No. 1.8e-30;
 Matches 342; Conservative 0; Mismatches 81; Indels 15; Gaps 3;
 QY 291 TGGCCGGGTGGTGGCTCATGCTCTAATCCAGCAGCTTTGGGAGGCGAAGGTGAGTAG 350
 Db |||||
 Db 436 TGGCAAGGCACAGTGGCTCAAGCTGTAAATCCAGCAGCTTTGGGAGGCGGAGGTGGTGG 377
 QY 351 ATCACTTCAGCGCGGAGTTTGGACACAGCCTGGCCACATGGTGAACCCCATCTGTAC 410
 Db |||||
 Db 376 ATCACTTCAGGTTCAGAGTTTGAAGACAGCCTGGCCACATGGTGAACCCCATCTGTAC 317
 QY 411 TAAATAACAAAAGTT-AGCTGGGGGTGGTGTAGATGCTGTAGTCCAGCTACTTGGG 469
 Db |||||
 Db 316 TAAATAACAAAANTGGCCGGGGTGGTGTAGTCAAGCTGTAAATCCAGCAGTTGGG 257
 QY 470 AGGCTGAGGATGAGAAATC-----GCTTGAGCCAGCCTGGGCAATACAGCAA 517
 Db |||||
 Db 256 AGGCTGAGGTGGTGGATCAAGGTCAGAGATCGAGACCATCTCTGCTAACAGATGA 197
 QY 518 GACCCGCTCTTACAAATAAATAACAAAATTTAGTGGATGCTGGTGCATGCTGTA 577
 Db |||||
 Db 196 AACCTGCTCTTACTA--AAAATAACAAAATTTAGCCGGTGGTGGTGGCGGCTGTA 139
 QY 578 GTCTTAGCTGTAGGAGGCTGAGATGGAAGATTGCTTGAGCCTGGGAGGTCAAGGCTG 637
 Db |||||
 Db 138 GTCCAGCTACTTGGAGGCTGAGCAGGAGATGGCGTGAATCTGGGAGGCGAGACTTG 79
 QY 638 CAGTGAGCGGAGATGGCGCACTGCACTCCAGCCTGGGCAACAGAGTGAGACCTGTCTC 697
 Db |||||
 Db 78 CAGTGAGCAGATCGTGCCTGCACTCCAGCCTGGGCGAGAGCGAGACTCTGTCTC 19
 QY 698 AGAAAAAATAAAAAA 715
 Db |||||
 Db 18 AAAAAAATAAAAAA 1

RESULT 9

CF887062
 LOCUS
 DEFINITION
 UI-CF-ECO-abg-d-18-18-UI-r18 UI-CF-ECO Homo sapiens cDNA clone
 UI-CF-ECO-abg-d-18-18-UI 5', mRNA sequence.
 CF887062
 ACCESSION
 VERSION
 KEYWORDS
 EST.
 SOURCE
 Homo sapiens (human)
 ORGANISM
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE
 1 (bases 1 to 680)
 AUTHORS
 Bonaldo,M.F., Lennon,G. and Soares,M.B.
 TITLE
 Normalization and subtraction: two approaches to facilitate gene
 discovery

JOURNAL
MEDLINE
PUBMED
COMMENT

Genome Res. 6 (9), 791-806 (1996)
 97044477
 889548
 Contact: McCray, PB
 McCray Lab
 University of Iowa
 204 University of Iowa Med Labs, Iowa City, IA 52242, USA
 Tel: 319 356 4866
 Fax: 319 356 7171
 Email: paul-mccray@uiowa.edu
 Tissue Procurement: Dr. M. J. Welsh, University of Iowa
 cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: Distribution information can be found at
 http://genome.uiowa.edu/distribution/cf.html
 The following repetitive elements were found in this cDNA
 sequence: 176-371 >ALU 309-581, >ALU 605-678, >ALU
 Seq primer: M13 REVERSE.

FEATURES
source

Location/Qualifiers
 1..680
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="UI-CF-ECO-abg-d-18-18-UI"
 /tissue_type="Lung"
 /dev_stage="Adult and Fetal"
 /lab_host="DH10B (Life Technologies) (T1 phage resistant)"
 /clone_lib="UI-CF-ECO"
 /note="Organ: Lung; Vector: pT73-Pac (Pharmacia) with a
 modified polylinker; Site 1: EcoR I; Site 2: Not I;
 UI-CF-ECO is a cDNA library containing the following
 tissue(s): Normal lung from adult and from fetal day 64,
 day 87, week 19 and week 42. The library was constructed
 according to Bonaldo, Lennon and Soares, Genome Research,
 6:791-806, 1996. First strand cDNA synthesis was primed
 with an oligo-dT primer containing a Not I site. Double
 stranded cDNA was ligated to an EcoR I adaptor, digested
 with Not I, and cloned directionally into pT73-Pac
 vector. The oligonucleotide used to prime the synthesis of
 first-strand cDNA contains a library tag sequence that is
 located between the Not I site and the (dT)18 tail. The
 sequence tag for this library is AGTGCCTTAC."

ORIGIN

Query Match 12.3%; Score 247.6; DB 7; Length 680;
 Best Local Similarity 78.5%; Pred. No. 2.4e-30;
 Matches 335; Conservative 0; Mismatches 85; Indels 7; Gaps 3;
 QY 293 GCCGGTGTGGTGGCTCATGCTCTAATCCAGCAGCTTTGGGAGGCGAAGGTGAGTAGT 352
 Db |||||
 Db 176 GCTGGCAGAGTGGTGTGCTGTGCTGTAACTCAGCAGCTTT-AGAGGCCAAGCGAGCGAGAC 234
 QY 353 CACTTGGGCGGGAGTTTGAGCAGCCTGGCCAACTGGTGAACCCCATCTGTACTA 412
 Db |||||
 Db 235 CTCCTTGAGCCAGGAGTTTGAGCAGCCTGGCCAACTGGTGAACCCCATCTGTACTA 294
 QY 413 AAAATACAAAAGTTAGCTGGCGTGGTGTAGATGCTGTAGTCCAGCTACTTGGGAGG 472
 Db |||||
 Db 295 AAAACACAAAATAGCGCGGCATGTGGCTCAGCCCTGTAAATCCAGCAGCTTTGGGAGG 354
 QY 473 CTGAGGCTAGAGATCGCTTGGCCAGCCTGGGCAATACAGCAAGACCCGCTCTCTACA 532
 Db |||||
 Db 355 CTGAGACCGAGGAGATC----GAGACCATTTCTGGCTTAACACTGTGAACCCCGTCTCTACT 410
 QY 533 AATAAATAACAAAAAATTAGTTGGATGTGGTGCATGCTCCCTGTAGTCTTACTGTAGG 592
 Db |||||
 Db 411 A--AAAAATACAGAAATTAGCTGGGTATGGTGGCGCTGTAGTCCAGCTACTCGG 468
 QY 593 GAGGCTGAGATGGAAGGATTGCTTGAGCCTGGGAGGTCAAGGCTCAAGTGCAGTGCAGATG 652
 Db |||||
 Db 469 GAGGCTGAGGAGGAGAGTGGCATGAACCCCGGAGGAGAGCTTGCAGTGCAGTGCAGATC 528

QY 653 GCGCCACTCCAGCCTGGCAACAGTAGTGAGACCCCTGTCTCAGAAAAA 712
 |||||
 Db 529 ACACCACTGCCTCCAGCCTGGCGACAGGCTGAGACTCGTCTCANAACAAACAAA 588
 |||||
 QY 713 AAAAAA 719
 |||||
 Db 589 AACAAAA 595
 |||||

BU677785 699 bp mRNA linear EST 07-OCT-2002
 UI-CF-ECO-abg-d-18-0-UI-s1 UI-CF-ECO Homo sapiens cDNA clone
 UI-CF-ECO-abg-d-18-0-UI 3', mRNA sequence.
 BU677785
 BU677785.1 GI:23524101
 EST.
 Homo sapiens (human)
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 699)
 Bonaldo,M.F., Lennon,G. and Soares,M.B.
 Normalization and subtraction: two approaches to facilitate gene
 discovery
 Genome Res. 6 (9), 791-806 (1996)
 97044477
 889548
 Contact: McCray, PB
 McCray Lab
 University of Iowa
 2024 University of Iowa Med Labs, Iowa City, IA 52242, USA
 Tel: 319 356 4866
 Fax: 319 356 7171
 Email: paul-mccray@uiowa.edu
 Tissue Procurement: Dr. M. J. Welsh, University of Iowa
 cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
 DNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: Researchers may obtain clones from Research
 Genetics (www.resgen.com) or from Open Biosystems
 (www.openbiosystems.com).
 The following repetitive elements were found in this cDNA
 sequence: 9-182, >ALU (matched complement) 206-478, >ALU (matched
 complement) 416-611, >ALU (matched complement)
 Seq primer: M13 FORWARD
 POLYA=Yes.

FEATURES
 source
 1..699
 Location/Qualifiers
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="UI-CF-ECO-abg-d-18-0-UI"
 /tissue_type="Lung"
 /dev_stage="Adult and Fetal"
 /lab_host="DH10B (Life Technologies) (T1 phage resistant)"
 /clone_lib="UI-CF-ECO"
 /note="Organ: Lung; Vector: pT7T3-Pac (Pharmacia) with a
 modified polylinker; Site_1: EcoR I; Site_2: Not I;
 UI-CF-ECO is a cDNA library containing the following
 tissue(s): Normal lung from adult and from fetal day
 day 87, week 19 and week 42. The library was constructed
 according to Bonaldo, Lennon and Soares, Genome Research,
 6:791-806, 1996. First strand cDNA synthesis was primed
 with an oligo-dT primer containing a Not I site. Double
 stranded cDNA was ligated to an EcoR I adaptor, digested
 with Not I, and cloned directionally into pT7T3-Pac
 vector. The oligonucleotide used to prime the synthesis of
 first-strand cDNA contains a library tag sequence that is
 located between the Not I site and the (dfr)18 tail. The
 sequence tag for this library is AAGTGTCTAC.
 TAG_TISSUE=Normal Lung Epithelial Cells Tissue nos 369-371
 and 380-383

ORIGIN

Query Match 12.3%; Score 247; DB 5; Length 699;
 Best Local Similarity 78.5%; Pred. No. 3e-30;
 Matches 335; Conservative 0; Mismatches 85; Indels 7; Gaps 3;
 TAG_LIB=UI-CF-ECO
 TAG_SEQ=AAGTGTCTAC

QY 293 GCGGGGTGGTGGCTCATGCTTAAATCCAGCACTTTGGAGGCGCAAGGTGAGTAGAT 352
 |||||
 Db 611 GTTGGGCACAGTGGCTCATGCTTAAATCTCAGCACTTT-AGAGGCGCAAGGCGAGAG 553
 |||||
 QY 353 CACTTGAGCCCGGAGTGGTGGAGACCAAGCTGGCCCAACATGGTGAACCCCATCTGTACTA 412
 |||||
 Db 552 CTCCTTGAGCCGAGGAGTTTGAGACCAAGCTGGCCCAAGCTGGCGAAGCCCATCTTACTA 493
 |||||
 QY 413 AAAATACAAAAGTTAGCTGGGCGTGGTGGTAGATGCTGTAGTCCAGCACTTCTGGGAGG 472
 |||||
 Db 492 AAAACACAAAAATAGCGCGGCATGGTGGCTCAGGCTGTAAATCCAGCACTTTGGGAGG 433
 |||||
 QY 473 CTGAGGCATGAGANTCGTTGAGCCAGCTGGGCGAATACAGCAAGACCCGCTCTTACTA 532
 |||||
 Db 432 CTGAGCCAGGAGATC---GAGACCAATCTGGCTAACACACTGTGAACCCCGTCTTACT 377
 |||||
 QY 533 AATAAAATACAAAAATAGTTGGATGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 592
 |||||
 Db 376 A-AAATACAAAGAAATTAAGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 319
 |||||
 QY 593 GAGGCTGAGATGGAAGATTGCTTGGAGCTGGGAGGTCAAGGCTGAGTGAGCGGAGATG 652
 |||||
 Db 318 GAGGCTGAGGAGGAGATGGCATGAACCCGAGGAGAGCTTGCAGTGAGCTGAGATC 259
 |||||
 QY 653 GCGCCACTGCCTCCAGCCTGGCAACAGTAGTGAGACCCCTGTCTCAGAAAAA 712
 |||||
 Db 258 ACACCACTGCCTCCAGCCTGGCGACAGGCTGAGACTCCGTTCAAAAAACAAAAACAAA 139
 |||||
 QY 713 AAAAAA 719
 |||||
 Db 198 AACAAAA 192
 |||||

RESULT 11

CF890780/c
 LOCUS
 DEFINITION
 UI-CF-ECO-abg-d-18-18-UI-s18 UI-CF-ECO Homo sapiens cDNA clone
 UI-CF-ECO-abg-d-18-18-UI 3', mRNA sequence.
 CF890780
 CF890780.1 GI:38149850
 EST.
 SOURCE
 Homo sapiens (human)
 ORGANISM
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 753)
 Bonaldo,M.F., Lennon,G. and Soares,M.B.
 Normalization and subtraction: two approaches to facilitate gene
 discovery
 Genome Res. 6 (9), 791-806 (1996)
 97044477
 889548
 Contact: McCray, PB
 McCray Lab
 University of Iowa
 2024 University of Iowa Med Labs, Iowa City, IA 52242, USA
 Tel: 319 356 4866
 Fax: 319 356 7171
 Email: paul-mccray@uiowa.edu
 Tissue Procurement: Dr. M. J. Welsh, University of Iowa
 cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
 DNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: Researchers may obtain clones from Research
 Genetics (www.resgen.com) or from Open Biosystems
 (www.openbiosystems.com).
 The following repetitive elements were found in this cDNA
 sequence: 9-182, >ALU (matched complement) 206-478, >ALU (matched
 complement) 416-611, >ALU (matched complement)
 Seq primer: M13 FORWARD
 POLYA=Yes.

REFERENCE
 AUTHORS
 TITLE
 Normalization and subtraction: two approaches to facilitate gene
 discovery
 Genome Res. 6 (9), 791-806 (1996)
 MEDLINE
 PUBMED
 COMMENT


```

Db      507 CTACTCAGAGGCTGAGCGAGGAGATGCGCGAACCCTGGAGCGGAGCTTGCACTGAG 566
QY      645 CCGAGATGGCGCCACTGCCTCCAGCTGGGCAACAGAGTGAGACCCCTGCTCAGAA--- 701
Db      567 CCAAGATCGGCCACTGCACTCCAGCTGGTGCACAGAGTGAGACTCCATCTCNAAAAC 626
QY      702 AAAAAAAAAAAAAAAAAAGGAGGAGAGAGACTCAAGCACGCCCTCCACAGGACTGCT 761
Db      627 AAAAAACAAAAACAAAAACAGCTAAGACTTTCAACTAAGCATCANGATGGGAGGAG 686
QY      762 GAGGCCCTGCAGTGTCTGCAGCATGTGGCCCGAGCGCGGAGACTCTGTAAAGCC 815
Db      687 GAGGTGTG-GGGTTTTTGAAGGAAGTGACCTCAAGCTGAGCACCCCTGTTTCCC 739

RESULT 13
BC039515
LOCUS   BC039515      838 bp      mRNA      linear      HTC 04-MAR-2003
DEFINITION Homo sapiens, clone IMAGE:5585187, mRNA.
ACCESSION BC039515
VERSION   BC039515.1 GI:24659995
KEYWORDS
SOURCE   Homo sapiens (human)
ORGANISM Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
          1 (bases 1 to 838)
          Strausberg, R.
          Direct Submission
          Submitted (01-NOV-2002) National Institutes of Health, Mammalian
          Gene Collection (MGC), Cancer Genomics Office, National Cancer
          Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
          USA
          NIH-MGC Project URL: http://mgc.nci.nih.gov
          Contact: MGC help desk
          Email: cgabbs-remail.nih.gov
          Tissue Procurement: ATCC
          CDNA Library Preparation: Life Technologies, Inc.
          CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
          DNA Sequencing by: Sequencing Group at the Stanford Human Genome
          Center, Stanford University School of Medicine, Stanford, CA 94305
          Web site: http://www.shgc.stanford.edu
          Contact: (Dickson, Mark) mcd@paxil.stanford.edu
          Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
          R. M.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAX Plate: 88 Row: h Column: 14
This clone has the following problem: retained intron.

FEATURES
source
    1..838
    /organism="Homo sapiens"
    /mol_type="mRNA"
    /db_xref="taxon:9606"
    /clone="IMAGE:5585187"
    /tissue_type="Testis, embryonal carcinoma"
    /clone_lib="NIH_MGC_92"
    /lab_host="DH10B"
    /note="Vector: pCMV-SPORT6"

ORIGIN
Query Match      12.3%; Score 246.6; DB 3; Length 838;
Best Local Similarity 74.1%; Pred. No. 3.3e-30;
Matches 358; Conservative 0; Mismatches 109; Indels 16; Gaps 3;

QY      239 TTGGCCCTACAAGTTGTGTGATGGGTACCCATGCGCAAGTGTGAGGGAGGCTGCGCGG 298
Db      168 TGTGCTCTGAAATCCTCAGCTGATACTTAGTTTAAATAGTCTCAAGTAGGCTGGG 227
QY      299 TGTGTTGGCTATGCTCTTAATCCAGCACTTTGGAGGCCCAAGGTGATAGATCACTTG 358
Db      228 CGTGGTGGCTCACACCTGTGAATCCCACTTTGGAGGCCCGAGGCGGCGATCACCTG 287

```

```

QY      359 AGCGCGGAGTTTGACACACAGCTGGCCAAACATGGTGAACCCCATCTGTACTAAAAATA 418
Db      288 AGGTACAGGAGTTTCAAGACACAGCTGGCCAAACATGGTGAACCCCATCTGTACTAAAAATA 347
QY      419 CAAAGTTAGCTGGGGCTGGTGTAGATCCCTGTAGTCCAGCTACTTTGGGAGGCTGAGG 478
Db      348 CAAAA--TTAGCTGGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 406
QY      479 CATGAGAATC-----GCTTGAGCCCGAGCTGGGCAATACAGCAAGACCCCGTC 526
Db      407 CAGGTGGATCAGAGGCTCAGGGGTTTCGAGACCGCTGACCAACATGGTGAACCCCTGTC 466
QY      527 TCTCAAAATAAAATACAAAAATTTAGTTGGATGTGGTGGTGGTGGTGGTGGTGGTGGTGG 586
Db      467 TCTAC---TAAAAATACAAAAATTTAGCCGGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 523
QY      587 GCTAGGAGGCTGAGATGAAGGATTGCTTTGAGCCTGGGAGGTCAAGGCTGCAGTGCAGCC 646
Db      524 ACTCGGAGGCTGAAGCGGGAGAAATTGTTGAACGGGGAGGGAGGTTGCAGTGAACC 583
QY      647 GAGATGGCGCCACTGCCTCCAGCTGGGCAACAGAGTGAGACCCCTGTCTCAGAAAAAAA 706
Db      584 AGGATCGCGCCACTGCCTCCAGCTGGGCGACAGAGCGAGACTCCGTCTCAAAAAAAA 643
QY      707 AAA 709
Db      644 AAA 646

BM469250      1060 bp      mRNA      linear      EST 05-FEB-2002
AGENCOURT_6444645 NIH_MGC_92 Homo sapiens cDNA clone IMAGE:5585187
5', mRNA sequence.

ACCESSION   BM469250
VERSION     BM469250.1 GI:18518292
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
          1 (bases 1 to 1060)
          NIH-MGC http://mgc.nci.nih.gov/.
          National Institutes of Health, Mammalian Gene Collection (MGC)
          Unpublished (1999)
          Contact: Robert Strausberg, Ph.D.
          Email: cgabbs-remail.nih.gov
          Tissue Procurement: ATCC
          CDNA Library Preparation: Life Technologies, Inc.
          CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
          DNA Sequencing by: Agencourt Bioscience Corporation
          Clone distribution: MGC clone distribution information can be
          found through the I.M.A.G.E. Consortium/LLNL at:
          http://image.llnl.gov
          Plate: LLAM12351 row: b column: 04
          High quality sequence stop: 659.
          Location/Qualifiers
            1..1060
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="IMAGE:5585187"
            /tissue_type="embryonal carcinoma, cell line"
            /lab_host="DH10B (phage-resistant)"
            /clone_lib="NIH_MGC_92"
            /note="Organ: testis; Vector: pCMV-SPORT6; Site 1: NotI;
            Site 2: SalI; Cloned unidirectionally; oligo-dT primed.
            Average insert size 2.5 kb. Library enriched for
            full-length clones and constructed by Life Technologies.
            Note: this is a NIH_MGC Library."

FEATURES
source
    1..1060
    /organism="Homo sapiens"
    /mol_type="mRNA"
    /db_xref="taxon:9606"
    /clone="IMAGE:5585187"
    /tissue_type="embryonal carcinoma, cell line"
    /lab_host="DH10B (phage-resistant)"
    /clone_lib="NIH_MGC_92"
    /note="Organ: testis; Vector: pCMV-SPORT6; Site 1: NotI;
    Site 2: SalI; Cloned unidirectionally; oligo-dT primed.
    Average insert size 2.5 kb. Library enriched for
    full-length clones and constructed by Life Technologies.
    Note: this is a NIH_MGC Library."

ORIGIN

```

```
Query Match      12.3%; Score 246.6; DB 4; Length 1060;
Best Local Similarity 74.1%; Pred. No. 3.1e-30;
Matches 358; Conservative 0; Mismatches 109; Indels 16; Gaps 3;

QY 239 TTTGGCTACAAAGTTGTGGATGTGGSTRACCATGCCAAGTGTGAGGGAGAGCTGGCCGGG 298
DB 168 TGTGGTCTGGAATACTCCAGGCTGATAACTTAGTTTAAATAAGTGTCTCAAGTAGGCTGGG 227

QY 299 TGTGGTGTCTCATGCTCTTAATCCAGCACATTTGGAGGCCAAGGTGTAGTGTAGTACACATG 358
DB 228 CGTGTGTGTCTCACCTGTGTATCCACACATTTGGAGGCCGAGGAGCGGATCACCTG 287

QY 359 AGGCGGGAGTTTGAGACACAGCTGGCCAAACATGTGAAACCCCATCTGTACTAAAAATA 418
DB 288 AGGTCAAGAGTTCAAGACACAGCTGGCCAAACATGTGAAACCCCATCTGTACTAAAAATA 347

QY 419 CAAAAGTTAGCTGGCGTGTGTGTAGATGCTGTGTAGTCCAGCTACTTTGGAGGCTGAGG 478
DB 348 CAAAATTTAGCTGGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 406

QY 479 CATGAGATC-----GCTTGAGCCAGCTGGGCAATACAGCAAGACCCCGTC 526
DB 407 CAGGTGATCAGAGGTCAAGGCTGAGAGCCAGCTGACCAACATGTGTGAAACCCCTGTC 466

QY 527 TCTACAAATAAAATACAAAAAATTAGTTGGATGTGGTGTGTGTGTGTGTGTGTGTGTGTGTGT 586
DB 467 TCTAC---TAAATAACAAAAAATTAGCCGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 523

QY 587 GCTAGGAGGCTGAGATGGAAGATGCTGTGAGCTGTGGAGGTCAAGCTGCAAGTGAGCC 646
DB 524 ACTCGGAGGCTGAAGCGGGAGATTTGCTTGAACCGGGAGCGGAGGTTCAGTGAACC 583

QY 647 GAGATGGCGCCACTCACTCCAGCTGGGCAACAGAGTGAGACCTGTCTCAGAAAAAAA 706
DB 584 AGGATCGGCCCACTGCACTCCAGCTGGGCAACAGAGTGAGACCTGTCTCAGAAAAAAA 643

QY 707 AAA 709
DB 644 AAA 646
```

```
RESULT 15
LOCUS      CB963738
DEFINITION AGENCOURT 13464887 NIH MGC 187 Homo sapiens cDNA clone
IMAGE:30320300 5', mRNA sequence.
ACCESSION  CB963738
VERSION     EST.
KEYWORDS    Homo sapiens (human)
SOURCE      Homo sapiens
ORGANISM    Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 733)
NIH-MGC http://mgc.mci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Dr. Michael Brownstein and Dr. Miklos Palkovits
cDNA Library Preparation: CLONTECH Laboratories, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: NDCM117 row: n column: 21
High quality sequence stop: 550.
Location/Qualifiers
1..733
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
```

FEATURES
source

```
/clone="IMAGE:30320300"
/lab_host="DH10B (TI phage-resistant)"
/clone_lib="NIH MGC 187"
/note="Organ: Blood vessels - aorta, basilar and artery;
Vector: pNIR-LIB; Site 1: SfiI (ggccattatggcc); Site 2:
SfiI (ggccgctcgcc); 5' and 3' adaptors were used in
cloning as follows: 5' adaptor sequence:
5'-CACGCGCAATATGCGC-3' and 3' adaptor sequence:
5'-ATTAGAGCGCGGCGGCGACATG-DT(30)BN-3' (where B = A,
C, or G and N = A, C, G, or T). Average insert size 1.4 kb
(range 0.5-4.0 kb). 14/15 colonies contained inserts by
PCR. This library was enriched for full-length clones and
was constructed by Clontech Laboratories (Palo Alto, CA).
Note: this is a NIH_MGC Library."
```

ORIGIN

```
Query Match      12.3%; Score 246.4; DB 6; Length 733;
Best Local Similarity 76.9%; Pred. No. 3.7e-30;
Matches 332; Conservative 0; Mismatches 86; Indels 14; Gaps 2;

QY 294 CCGGCTGTGGTGTCTCATGCTCTTAATCCAGCACTTTGGAGGCCAAGGTGTAGTAGTC 353
DB 99 CTGGCGCGAGTGTTCACGCTGTAAATCCAGCACTTTGGAGGCCAAGGCGGTGTATC 158

QY 354 ACTTGAGCCCGGAGTTTGACACCAAGCTGCGCAACATGTGTGAAACCCCATCTGTACTAA 413
DB 159 ACCTGAGTTTGGAGTTTCGAGACCAAGCTGCGCAACATGTGTGAAACCCCATCTGTACTGA 218

QY 414 AAATACAAAGTTAGCTGGCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 473
DB 219 AAATACAAAGTTAGCTGGCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 278

QY 474 TGAGGCATGAGATC-----GCTTGAGCCCGCCTGGGCAATACAGCAAGACC 521
DB 279 TGAGCGGGTGTGATCAGGAGTTCAGAGATCGAGACCATCTCTGGCTAACATGTGTGAAACC 338

QY 522 CGCTCTTACAAATAAAATACAAAAAATTAGTTGGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 581
DB 339 CCATCTCTACTA--AAAAATACAAAAAATTATCTGGCGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 396

QY 582 TAGCTGTAGGAGGCTGAGATGGAAGGATTGCTTGAGCTGGGAGGTCAAGGCTGCAGT 641
DB 397 CAGCTTACTCAGAGGTTGAGGAGGAGATCGCTTGAACCCAGAGGAGGAGGTTGCAGT 456

QY 642 GAGCCGAGATGGCGCCACTGCACTCCAGCTGGGCAACAGAGTGAGAGCCCTGTCTCAGAA 701
DB 457 GAGCCGAGATTGGCGCCACTGCACTCCAGCTGGGTTGACAGAGCAAGACTCCACCTCAAAG 516

QY 702 AAAAAA 713
DB 517 TAATAATAATA 528
```

Search completed: November 23, 2004, 12:26:22
Job time : 5943.85 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 22, 2004, 15:08:20 ; Search time 890.868 Seconds
(without alignments)
11849.778 Million cell updates/sec

Title: US-10-017-724-3_COPY_85400_87410

Perfect score: 2011

Sequence: 1 ccggccgcttgaagagctt.....gaatgtggaagtaggcggt 2011

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4134886 seqs, 2624710521 residues

Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_23Sep04:*

1: Geneseqn1980s:*

2: Geneseqn1990s:*

3: Geneseqn2000s:*

4: Geneseqn2001as:*

5: Geneseqn2001bs:*

6: Geneseqn2002as:*

7: Geneseqn2002bs:*

8: Geneseqn2003as:*

9: Geneseqn2003bs:*

10: Geneseqn2003cs:*

11: Geneseqn2003ds:*

12: Geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	2011	100.0	98229	9	AAD55116 Human ACE
2	1233.6	61.3	2849	2	AAG35034 DNA fragm
3	824.6	41.0	1856	3	AAG38329 Human ang
4	824.6	41.0	1856	10	ADF39358 Intron 16
5	599.4	29.8	616	10	ADC26746 Human lip
6	270.4	13.4	41369	10	ADL13522
7	267.2	13.3	2836	11	ADM01345
8	267	13.3	12730	4	AAK70514
9	265.4	13.2	300000	10	ADE86352
10	265.4	13.2	300001	12	ADO14076
11	262.8	13.1	26874	8	AAK56105
12	262.8	13.1	26874	9	ADA02467
13	262.8	13.1	26874	10	ADB72206
14	262	13.0	20951	6	AAU53468
15	261.4	13.0	2568	6	AAK99739
16	260.8	13.0	3769	4	AAK67255
17	259.2	12.9	1496	6	AAU50885
18	258.6	12.9	3276	5	AAU50098
19	258.6	12.9	3276	10	ADB33435
20	257.8	12.8	4301	4	AAI68146
21	257.8	12.8	4301	10	ADL13487

C 22	257.8	12.8	4303	3	AAA95826	Aaa95826 Human met
C 23	257.8	12.8	4307	10	ADB85504	Adb85504 Human egg
C 24	257.8	12.8	4307	10	ACF63399	Act63399 Human ADA
C 25	257.8	12.8	4406	4	AAS46100	Aas46100 Human DNA
C 26	257.8	12.8	4407	3	AAA37111	Aaa37111 Human PRO
C 27	257.8	12.8	4407	4	AAF54426	Aaf54426 Probe #46
C 28	257.8	12.8	4407	8	ACA89550	Acg89550 cDNA enco
C 29	257.8	12.8	4407	8	ACA73560	Acg73560 Human sec
C 30	257.8	12.8	4407	8	ACA05875	Aca05875 Human sec
C 31	257.8	12.8	4407	8	ACA66709	Act66709 cDNA enco
C 32	257.8	12.8	4407	8	ACF20284	Acf20284 Human sec
C 33	257.8	12.8	4407	8	ACF19670	Acf19670 Human sec
C 34	257.8	12.8	4407	8	ACD21958	Acd21958 Human sec
C 35	257.8	12.8	4407	8	ACF13123	Acf13123 Human sec
C 36	257.8	12.8	4407	8	ACD25226	Acd25226 Human sec
C 37	257.8	12.8	4407	8	ACF00275	Acf00275 Human sec
C 38	257.8	12.8	4407	8	ACA72332	Aca72332 Novel hum
C 39	257.8	12.8	4407	8	ACD04856	Acd04856 Novel hum
C 40	257.8	12.8	4407	8	ACD18317	Acd18317 Human sec
C 41	257.8	12.8	4407	8	ACD08324	Acd08324 Human sec
C 42	257.8	12.8	4407	8	ACA88758	Aca88758 Novel hum
C 43	257.8	12.8	4407	8	ACA70200	Aca70200 Human sec
C 44	257.8	12.8	4407	8	ACD12422	Acd12422 Novel hum
C 45	257.8	12.8	4407	8	ACC74337	Acc74337 Human sec

ALIGNMENTS

RESULT 1

AAD55116

ID AAD55116 standard; DNA; 98829 BP.

XX

AC AAD55116;

XX

DT 07-AUG-2003 (first entry)

XX

DE Human ACE reference DNA (GI 13027555).

XX

KW Thrombospondin 2; THBS2; angiotensin converting enzyme; polymorphism;

KW ACE-1; beta-fibrinogen; FGB; peripheral vascular disease; ischaemia;

KW vascular disease; myocardial infarction; pulmonary embolism; stroke;

KW atherosclerosis; coronary artery disease; venous thromboembolism; human;

KW gene; ds.

XX

OS Homo sapiens.

XX

PN WO2003020118-A2.

XX

PD 13-MAR-2003.

XX

PF 04-SEP-2002; 2002WO-US028113.

 XX |

CC involves contacting the nucleic acid with a complementary probe or
CC primer. The method is useful for diagnosing or aiding in the diagnosis of
CC vascular disease or disorder in a subject e.g. myocardial infarction,
CC coronary artery disease, atherosclerosis, ischemia, stroke, peripheral
CC vascular disease, venous thromboembolism and pulmonary embolism. The
CC present sequence is human ACE reference DNA. Note: This sequence is said
CC to encode SEQ ID NO: 4 (AAB36412). However this does not appear to be the
CC case
XX
SQ Sequence 98829 BP; 21920 A; 26068 C; 26526 G; 23318 T; 0 U; 997 Other;
Query Match 100.0%; Score 2011; DB 9; Length 98829;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2011; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CCGGCCCTTTGAAGAGACTACCCCGACACACAGGACCCGACACACATACCTCCGAGC 60
DB CCGGCCCTTTGAAGAGACTACCCCGACACACAGGACCCGACACATACCTCCGAGC 85459
QY 61 TCCCTCTCAACCCACCCCTTCCAGGTTGGAGAACTTGAGGCATAACTTGTCTCCATGA 120
DB TCCCTCTCAACCCACCCCTTCCAGGTTGGAGAACTTGAGGCATAACTTGTCTCCATGA 85519
QY 121 GGAATCTCAACCCAGAAATGGGTCTTCTGGCCCCCAGCCAGCTCCCAATTAGAA 180
DB GGAATCTCAACCCAGAAATGGGTCTTCTGGCCCCCAGCCAGCTCCCAATTAGAA 85579
QY 181 TGCAAAATAGAGGGGAATGGAATAAAGAGGAGAACGTTTCCAGGACAGGGTT 240
DB TGCAAAATAGAGGGGAATGGAATAAAGAGGAGAACGTTTCCAGGACAGGGTT 85639
QY 241 TGGCCTCAAGTTGTGGATGTGGGTACCCATGTCAGGAGGAGGCTGGCCGGGTG 300
DB TGGCCTCAAGTTGTGGATGTGGGTACCCATGTCAGGAGGAGGCTGGCCGGGTG 85699
QY 301 TGGTGGCTCATGCTCTAATCCAGCACTTTGGGAGGCCAAGGTGAGTAGATCACTTGAG 360
DB TGGTGGCTCATGCTCTAATCCAGCACTTTGGGAGGCCAAGGTGAGTAGATCACTTGAG 85759
QY 361 CGCGGGAGTTGAGACCGAGCTGGCCCAACATGGTGAACCCCATCTGTACTAAATAACA 420
DB CGCGGGAGTTGAGACCGAGCTGGCCCAACATGGTGAACCCCATCTGTACTAAATAACA 85819
QY 421 AAAAGTTAGTGGCGTGTGGTAGATGCTGTAGTCCAGCTACTTGGGAGGCTGAGGCA 480
DB AAAAGTTAGTGGCGTGTGGTAGATGCTGTAGTCCAGCTACTTGGGAGGCTGAGGCA 85879
QY 481 TGAGAAATCGCTTGAGCCAGCCCTGGGGAATACAGCAAGACCCCTCTCTCAAAATAAAT 540
DB TGAGAAATCGCTTGAGCCAGCCCTGGGGAATACAGCAAGACCCCTCTCTCAAAATAAAT 85939
QY 541 ACAGAAATATAGTTGGATGTGGTGTGCATGCTGTAGTCTGTAGTCTAGGAGGCTGA 600
DB ACAGAAATATAGTTGGATGTGGTGTGCATGCTGTAGTCTGTAGTCTAGGAGGCTGA 85999
QY 601 GATGAAGGATGTGTGAGCTGGAGGTCAAGGCTGAGTGCAGCCGAGATGGCCACT 660
DB GATGAAGGATGTGTGAGCTGGAGGTCAAGGCTGAGTGCAGCCGAGATGGCCACT 86059
QY 661 GCATCTCAGCTGGGCAACAGAGTGAGACCTGTCTCAGAAAAAATAAAAAA 720
DB GCATCTCAGCTGGGCAACAGAGTGAGACCTGTCTCAGAAAAAATAAAAAA 86119
QY 721 GGAGAGGAGAGAGACTCAAGCAGCCCTCAGAGACTGCTGAGGCCCTGCAAGGTGTCTG 780
DB GGAGAGGAGAGAGACTCAAGCAGCCCTCAGAGACTGCTGAGGCCCTGCAAGGTGTCTG 86179
QY 781 CAGCATGTGGCCCGAGCCGGGACTCTGTAAAGCACTGCTGGAGAGCCACTCCCATCT 840
DB CAGCATGTGGCCCGAGCCGGGACTCTGTAAAGCACTGCTGGAGAGCCACTCCCATCT 86239
QY 841 TTCTCCCATTTCTTAGACTGCTGCTCCATACAGTCACTTTTATGTGGTTTGGCAATTT 900
DB TTCTCCCATTTCTTAGACTGCTGCTCCATACAGTCACTTTTATGTGGTTTGGCAATTT 87379

DB 86240 TTCTCCCATTTCTTAGACTGCTGCTCCATACAGTCACTTTTATGTGGTTTGGCAATTT 86299
QY 901 TATTCAGCTCTGAAATTTCTCTGAGCTCCCTTAAAGCAGAGGTGAGCTAAGGGCTGGA 960
DB TATTCAGCTCTGAAATTTCTCTGAGCTCCCTTAAAGCAGAGGTGAGCTAAGGGCTGGA 86359
QY 961 GCTCAAGGCATTTAAACCCCTTACCAGATCTGACGAATGTGATGTCACCGTCCCGAAATA 1020
DB GCTCAAGGCATTTAAACCCCTTACCAGATCTGACGAATGTGATGTCACCGTCCCGAAATA 86419
QY 1021 TGAAGACTTTATGGGATGGAGGCTGCGGAGACAGGCGGGGAGAGCATCTCTCA 1080
DB TGAAGACTTTATGGGATGGAGGCTGCGGAGACAGGCGGGGAGAGCATCTCTCA 86479
QY 1081 GTTTTACCCGAAATACGTGGAACTCATCAACAGGCTGCCGGCTCAATGTGAGTCCCT 1140
DB GTTTTACCCGAAATACGTGGAACTCATCAACAGGCTGCCGGCTCAATGTGAGTCCCT 86539
QY 1141 GCTGCCAATCATCTGCGACCTTGGGTCCCTTCAATTTCTCAAGAGGTGCTGTGAAC 1200
DB GCTGCCAATCATCTGCGACCTTGGGTCCCTTCAATTTCTCAAGAGGTGCTGTGAAC 86599
QY 1201 CCAAGCTTAGAAAGGTAGATCCCTGAGAGGAGCAGTATGTGCTTGGGAGAGCT 1260
DB CCAAGCTTAGAAAGGTAGATCCCTGAGAGGAGCAGTATGTGCTTGGGAGAGCT 86659
QY 1261 GGCTGTGCTCCTCTGTAGGCTATGTAGATGAGGAGACTCGTGAGGCTTATGTACGAG 1320
DB GGCTGTGCTCCTCTGTAGGCTATGTAGATGAGGAGACTCGTGAGGCTTATGTACGAG 86719
QY 1321 ACACCATCTTGGAGCAAGACTGAGGCTCTTTCAGAGGCTGACGCTCTACCTC 1380
DB ACACCATCTTGGAGCAAGACTGAGGCTCTTTCAGAGGCTGACGCTCTACCTC 86779
QY 1381 AACCTGCATGCTAGTGGCGGCGCTGACCGTCACTACGCGGCCAGCACATCAAC 1440
DB AACCTGCATGCTAGTGGCGGCGCTGACCGTCACTACGCGGCCAGCACATCAAC 86839
QY 1441 CTGAGGCGGCCATTTCTGCTCACCTGTGGTAAAGGCACATGTCGGGCTTGAAGAG 1500
DB CTGAGGCGGCCATTTCTGCTCACCTGTGGTAAAGGCACATGTCGGGCTTGAAGAG 86899
QY 1501 GTAAAGACGACCAAGTGTGAGGTGGGACAGGCTGACTAGAGGTGAGGAGC 1560
DB GTAAAGACGACCAAGTGTGAGGTGGGACAGGCTGACTAGAGGTGAGGAGC 86959
QY 1561 AGGCTGGGACCTGAGAGACTCCAGCCCTGTGGGGATGTTGCCAGGCTGGAGGCGGT 1620
DB AGGCTGGGACCTGAGAGACTCCAGCCCTGTGGGGATGTTGCCAGGCTGGAGGCGGT 87019
QY 1621 GGGGCTGGGAGTGGGAGCCCCCCTTGCATCTGTGTCACATTCATCTGAGATCTAT 1680
DB GGGGCTGGGAGTGGGAGCCCCCCTTGCATCTGTGTCACATTCATCTGAGATCTAT 87079
QY 1681 GTCGGGCAAGTCAACATGGATGGGGGAAGAGTTAATCTTGTTCAGGAGACCAAGC 1740
DB GTCGGGCAAGTCAACATGGATGGGGGAAGAGTTAATCTTGTTCAGGAGACCAAGC 87139
QY 1741 ACCCATCAACATTTGTGATCTTAGAGGCGGAGAGAGGCTGTGATGGGAGCTGG 1800
DB ACCCATCAACATTTGTGATCTTAGAGGCGGAGAGAGGCTGTGATGGGAGCTGG 87199
QY 1801 GAGGCTTTTGGCAAGAGGTGGCTGTGAGCAGGCGCTCGGAAGATGACAGGTTTACAGA 1860
DB GAGGCTTTTGGCAAGAGGTGGCTGTGAGCAGGCGCTCGGAAGATGACAGGTTTACAGA 87259
QY 1861 TGGGAAGTGGGGGATGAGAGGACAGCGAGTGTTCAGGCCAGGGAATGGAACAAGA 1920
DB TGGGAAGTGGGGGATGAGAGGACAGCGAGTGTTCAGGCCAGGGAATGGAACAAGA 87319
QY 1921 AGAAGCTGAGAAATGTAATCTACTTCAACCTTGGAGCCCTCTTTTGGCAAGGCTGCAATC 1980
DB AGAAGCTGAGAAATGTAATCTACTTCAACCTTGGAGCCCTCTTTTGGCAAGGCTGCAATC 87379

QY	1981	TCAGATGCCCTGAATGCTGTGAAGTAGGCGGT	2011
DB	87380	TCAGATGCCCTGAATGCTGTGAAGTAGGCGGT	87410
RESULT 2			
AAQ35034			
ID	AAQ35034 standard; DNA; 2649 BP.		
XX	AAQ35034;		
XX	AC		
XX	25-MAR-2003	(revised)	
DT	27-MAY-1993	(first entry)	
XX			
XX	DNA fragment contg. ACE gene intron 16.		
XX			
KW	Angiotensin converting enzyme; polymorphism; detection; insertion;		
KW	vasoactive peptide metabolism; disease; hypertension; diagnosis;		
KW	atherosclerosis; granulomatous; sugar diabetes complications; ss.		
XX			
OS	Homo sapiens.		
XX			
FX	Key	Location/Qualifiers	
FT	intron	1..82	
FT		/*tag= a	
FT		/note= "end of intron 15"	
FT	exon	83..170	
FT		/*tag= b	
FT		/note= "exon 16"	
FT	intron	171..2026	
FT		/*tag= c	
FT		/note= "intron 16"	
FT	misc_feature	1621..1908	
FT		/*tag= f	
FT		/note= "repetitive Alu-type insertion sequence"	
FT	exon	2027..2170	
FT		/*tag= d	
FT		/note= "exon 17"	
FT	intron	2171..2649	
FT		/*tag= e	
FT		/note= "intron 17"	
XX			
XX	W09300360-A1.		
PN			
XX			
XX	07-JAN-1993.		
XX			
XX	24-JUN-1992;	92WO-FR000574.	
XX			
XX	27-JUN-1991;	91FR-00008020.	
XX			
XX	(INRM) INSERM INST NAT SANTE & RECH MED.		
PA			
PI	Soubrier F, Hubert C, Corvol P;		
XX			
XX	WPI; 1993-036331/04.		
DR			
XX			
PT	Detection of polymorphism in the angiotensin-I converting enzyme gene -		
PT	using new DNA sequences, for diagnosing atherosclerosis, granuloma-		
PT	associated diseases and diabetic complications.		
XX			
PS	Claim 1; Fig 2; 22pp; French.		
XX			
CC	The sequence is that of a DNA fragment contg. intron 16 of the gene for		
CC	angiotensin-I converting enzyme (ACE). Primers based on part of the		
CC	sequence of the intron can be used to detect polymorphism in the ACE		
CC	gene, esp. by hybridisation or amplification (partic. PCR), and to detect		
CC	the presence/absence of a defined insertion (i.e. nucleotides 1451-1738		
CC	of the intron sequence). This can be used to study the impact of		
CC	polymorphism on diseases, metabolism of vasoactive peptides and		
CC	mechanisms of hypertension. In particular it can be used in diagnosis of		
CC	atherosclerosis, granulomatous diseases and the complications of sugar		
CC	diabetes. (Updated on 25-MAR-2003 to correct PN field.)		

Db 589 CCAGCCCGTTTGAAGAGCTCACCCCGCACAAAGGACCCGCACACAGATACCTCCAGC 648
 QY 61 TCCCTCTCAACCCACCTTTCCAGGGTTGAGAACTTGGAGCAATACTGCTTCCATGA 120
 Db 649 TCCCTCTCAACCCACCTTTCCAGGGTTGAGAACTTGGAGCAATACTTCTTCATGA 708
 QY 121 GGAATCTCCACCCAGAAATGGGTCTTTCTGGCCCCCAGCCAGCTCCACATTTAGAACAA 180
 Db 709 GGAATCTCCACCCAGAAATGGGTCTTTCTGGCCCCCAGCCAGCTCCACATTTAGAACAA 768
 QY 181 TGACAAATGAAGGGGAATGGAATAAACAAGAGAAACGGTTTCCAGGACAGGGTT 240
 Db 769 TGACAAATGAAGGGGAATGGAATAAACAAGAGAAACGGTTTCCAGGACAGGGTT 828
 QY 241 TGGCTTACAAAGTTGTGGATGTGGGTACCCATGCCATGCCAAGTGTGAGGGAGGCTGCCGGTG 300
 Db 829 TGGCTTACAAAGTTGTGGATGTGGGTACCCATGCCAAGTGTGAGGGAGGCTGCCGGTG 888
 QY 301 TGGTGGCTCATGCTCTAATCCAGCACTTTGGAGGCCAAGGTGAGTAGATCACTTGAG 360
 Db 889 TGGTGGCTCATG-CTCTAATCCAGCACTTTGGAGGCCAAGGTGAGTAGATCACTTGAG 947
 QY 361 GCCGGAGTTGAGACAGCTGGCCCAACATGCTGTAACCCCATCTGTACTAAAAATACA 420
 Db 948 GCCGGAGTTGAGACAGCTGGCCCAACATGCTGTAACCCCATCTGTACTAAAAATACA 1007
 QY 421 AAAGTTAGCTGGCGGTGGTGTAGATGCTGTAGTCCAGCTACTTGGAGGCTGAGGCA 480
 Db 1008 AAAGTTAGCTGGCGGTGGTGTAGATGCTGTAGTCCAGCTACTTGGAGGCTGAGGCA 1067
 QY 481 TGAGAAATCGCTTGGAGCCAGCTGGCAATACAGCAAGACCCCGTCTTACAAATAAAT 540
 Db 1068 TGAGAAATCGCTTGGAGCCAGCTGGCAATACAGCAAGACCCCGTCTTACAAATAAAT 1127
 QY 541 ACAGAAATTAAGTTGAGTGTGGTGTAGTGTGATGCTGTAGTCTTGTAGTGTGAGGCTGA 600
 Db 1128 ACAGAAATTAAGTTGAGTGTGGTGTAGTGTGATGCTGTAGTCTTGTAGTGTGAGGCTGA 1187
 QY 601 GATGAAGGATTGCTTGAGCTGGAGGTCAAGGCTGCAAGTGGAGGCTGAGGCTGAGGCA 660
 Db 1188 GATGAAGGATTGCTTGAGCTGGAGGTCAAGGCTGCAAGTGGAGGCTGAGGCTGAGGCA 1247
 QY 661 GCATCTCAGCTTGGCAACAGAGTGAAGCTTGTCTTCAGAAAAAATAAATAAATAA 720
 Db 1248 GCATCTCAGCTTGGCAACAGAGTGAAGCTTGTCTTCAGAAAAAATAAATAAATAA 1305
 QY 721 GGAGAGGAGAGACTCAAGCAGCCCTTCAAGAGCTGCTGAGGCTTGCAGGTTCTG 780
 Db 1306 GGAGAGGAGAGACTCAAGCAGCCCTTCAAGAGCTGCTGAGGCTTGCAGGTTCTG 1365
 QY 781 CAGCATGTGGCCCCCAGGCGGGACTCTGTAAGCCACTGCTGGAGGCACTCCCATCT 840
 Db 1366 CAGCATGTG-CCGAGGCGGGGACTCTGTAAGCCACTGCTGGAGGCACTCCCATCT 1422
 QY 841 TTCTCCCATTTCTAGACCTGCTGCTATACAGTCACTTTTATGTGTTTT 891
 Db 1423 TTCTCCCATTTCTAGACCTGCTGCTATACAGTCACTTTTATGTGTTTT 1473

RESULT 4

ADF39358

ID ADF39358 standard; DNA; 1856 BP.

XX AC

XX ADF39358;

XX DT

DT 12-FEB-2004 (first entry)

XX DE

DE Introns 16 sequence of human ACE gene.

XX KW

KW Diabetes mellitus; renin-angiotensin II system;

KW angiotensin-converting enzyme; ACE; antidiabetic; insulin; hypoglycaemia;

KW ACE genotype; intron 16; human; ds.

XX

OS Homo sapiens.

XX US2003158090-A1.

XX PD

PD 21-AUG-2003.

XX XX

XX 04-OCT-2002; 2002US-00195330.

XX XX

XX 23-JUL-2001; 2001US-0306859P.

XX XX

XX (PEDE/) PEDERSEN-BJERGAARD U.

XX (AGER/) AGERHOLM-LARSEN B.

XX (THOR/) THORSTEINSSON B.

XX (PRAM/) PRAMMING S.

XX XX

XX Pedersen-Bjergaard U, Agerholm-Larsen B, Thorsteinsson B;

XX Prammig S;

XX WPI; 2003-778014/73.

XX XX

XX Treating diabetes mellitus in an individual comprises administering an

XX inhibitor of the renin-angiotensin II system and an antidiabetic.

XX Claim 36; Page 13-14; 17pp; English.

XX XX

XX The present invention relates to a method of treating diabetes mellitus

XX in an individual. The method comprises administering an inhibitor of the

XX renin-angiotensin II system (e.g. angiotensin-converting enzyme (ACE))

XX and an antidiabetic (e.g. insulin). Also disclosed is a method of

XX diagnosing susceptibility of an individual to hypoglycaemia. The method

XX involves detecting within a tissue sample the genotype of the ACE gene

XX and correlating the genotype with the susceptibility to hypoglycaemia.

XX The methods are useful for treating diabetes mellitus and hypoglycaemia.

XX The present sequence represents intron 16 of the human ACE gene.

XX CC

XX Sequence 1856 BP; 403 A; 546 C; 483 G; 424 T; 0 U; 0 Other;

XX SQ

XX Query Match 41.0%; Score 824.6; DB 10; Length 1856;

XX Best Local Similarity 98.3%; Pred. No. 5.5e-203;

XX Matches 876; Conservative 0; Mismatches 9; Indels 6; Gaps 4;

XX

XX 1 CGGCGCCGTTTGAAGAGCTCACCCCGCACAAAGGACCCGCACAGATACCTCCAGC 60

XX 589 CCAGCCCGTTTGAAGAGCTCACCCCGCACAAAGGACCCGCACAGATACCTCCAGC 648

XX 61 TCCCTCTCAACCCACCTTTCCAGGGTTGGAGAACTTGGAGCAATACTTGTCTTCATGA 120

XX 649 TCCCTCTCAACCCACCTTTCCAGGGTTGGAGAACTTGGAGCAATACTTGTCTTCATGA 708

XX 121 GGAATCTCCACCCAGAAATGGGTCTTTCTGGCCCCCAGCCAGCTCCACATTTAGAACAA 180

XX 709 GGAATCTCCACCCAGAAATGGGTCTTTCTGGCCCCCAGCCAGCTCCACATTTAGAACAA 768

XX 181 TGACAAATGAAGGGGAATGGAATAAACAAGAGAAACGGTTTCCAGGACAGGGTT 240

XX 769 TGACAAATGAAGGGGAATGGAATAAACAAGAGAAACGGTTTCCAGGACAGGGTT 828

XX 241 TGGCTTACAAAGTTGTGGATGTGGGTACCCATGCCAAGTGTGAGGGAGGCTGCCGGTG 300

XX 829 TGGCTTACAAAGTTGTGGATGTGGGTACCCATGCCAAGTGTGAGGGAGGCTGCCGGTG 888

XX 301 TGGTGGCTCATGCTCTAATCCAGCACTTTGGAGGCCAAGGTGAGTAGATCACTTGAG 360

XX 889 TGGTGGCTCATG-CTCTAATCCAGCACTTTGGAGGCCAAGGTGAGTAGATCACTTGAG 947

XX 361 GCCGGAGTTGAGACAGCTGGCCCAACATGCTGTAACCCCATCTGTACTAAAAATACA 420

XX 948 GCCGGAGTTGAGACAGCTGGCCCAACATGCTGTAACCCCATCTGTACTAAAAATACA 1007

XX 421 AAAGTTAGCTGGCGGTGGTGTAGATGCTGTAGTCCAGCTACTTGGAGGCTGAGGCA 480

XX 1008 AAAGTTAGCTGGCGGTGGTGTAGATGCTGTAGTCCAGCTACTTGGAGGCTGAGGCA 1067

XX Db

XX OS Homo sapiens.
XX PN WO2003054166-A2.
XX PD 03-JUL-2003.
XX PF 19-DEC-2002; 2002WO-US041225.
XX PR 20-DEC-2001; 2001US-0342603P.
XX PA (INCY-) INCYTE GENOMICS INC.
XX PI Jones KA, Schafer A;
XX PPI; 2003-559141/52.
XX DR
XX KW Determining susceptibility of an individual to joint space narrowing,
XX PT osteophyte development and/or joint pain comprises identifying whether
XX PT the individual has at least one polymorphism in a polynucleotide encoding
XX PT a protein.
XX PS Disclosure; SEQ ID NO 54; 297pp; English.
XX CC The invention relates to a method of determining susceptibility of an
XX CC individual to joint space narrowing and/or osteophyte development and/or
XX CC joint pain comprising identifying whether the individual has at least one
XX CC polymorphism in a polynucleotide encoding at least one of the protein
XX CC listed in the specification. The methods, composition and agent are
XX CC useful for modulating the susceptibility of an individual to joint space
XX CC narrowing and/or osteophyte development and/or joint pain that is
XX CC associated with a disease, preferably osteoarthritis. The cell line and
XX CC the non-human animal are useful for screening for an agent for diagnosing
XX CC an individual having susceptibility to joint space narrowing and/or
XX CC osteophyte development and/or joint pain. This sequence corresponds to
XX CC the polynucleotide encoding a protein listed in the specification. (Note:
XX CC The sequence data for this patent did not form part of the printed
XX CC specification but was obtained in electronic format directly from WIPO at
XX CC ftp.wipo.int/pub/published_pct_sequences).
XX SQ Sequence 41369 BP; 9847 A; 10483 C; 11706 G; 9333 T; 0 U; 0 Other;
Query Match 13.4%; Score 270.4; DB 10; Length 41369;
Best Local Similarity 77.1%; Pred. No. 6.5e-59;
Matches 347; Conservative 0; Mismatches 91; Indels 12; Gaps 1;
QY 292 GCGCGGTTGGTGGCTCATGCTCTAATCCAGCACTTTGGAGGCCAAGTGAGTGA 351
DB 33439 GGTCAAGGATGGTGGCTCATGCTCTAATCCCAACACTTTGGAGGCTAAGGTGGGCGA 33498
QY 352 TCACCTTGAGGCGGGAGTTTGAGACGAGCTCGCCCAACATGTGAACCCCATCTGTACT 411
DB 33499 TCACCTGAGTCAAGAGTTTGAGACGAGCTCGCCCAACATGTGAACCCCATCTGTACT 33558
QY 412 AAAAATACAAAGTTAGTGGCGGTGGTGGTGTAGATGCCCTGTAGTCCCACTACTTTGGAG 471
DB 33559 AAAAATACAAAGTTAGTGGCGGTGGTGGTGTAGATGCCCTGTAGTCCCACTACTTTGGAG 33618
QY 472 GCTGAGGCTATGAGATC-----GCTTGAGCCAGCTCGGCGCAATACAGACAGA 519
DB 33619 GCGGAGGCGGGGATACAGAGATCAGGAGATTTGAGACCATCTCGGCTAACACAGTGAAA 33678
QY 520 CCCCCTCTCTACAAATAAAATACAAATAATTTGATGTTGGATGTGGTGCATGCCCTGTAGT 579
DB 33679 CCTGTCTCTACAAATAAAATACAAATAATTTGATGTTGGATGTGGTGCATGCCCTGTAGT 33738
QY 580 CCTAGCTGCTAGGAGGCTGAGATGGAAGGATTTGCTTGAGCTCGGAGGCTCAAGCTGCA 639
DB 33739 CCGAGCTATTCCGAGGCTGAGCAGGAGATGGGTGNAACCCGAGAGGAGGAGCTTGA 33798
QY 640 GTGAGCCAGATGGCGCACTGCACTCCAGCTCGGCGCAACAGAGTGAGACCCCTGTCTCAG 699
DB 33799 GTGAGCCAGATGGCGCACTGCACTCCAGCTCGGCGCAACAGAGTGAGACCCCTGTCTCAG 33858

QY 700 AAAAAAAAAAAAAAAAAAAAAAAAAAGGAGGAG 729
DB 33859 AAAAAATATAAAATATAAAATATAAAATAG 33888
RESULT 7
ADM01345/C
ID ADM01345 standard; cDNA; 2836 BP.
XX
XX ADM01345;
XX
XX 20-MAY-2004 (first entry)
XX
XX Human cDNA of the invention SEQ ID NO:30.
XX
XX ss; gene; human; gene therapy; diagnostic marker; pharmaceutical.
XX
XX Homo sapiens.
XX
XX EP1347046-A1.
XX
XX 24-SEP-2003.
XX
XX 12-APR-2002; 2002EP-00008400.
XX
XX 22-MAR-2002; 2002JP-00137785.
XX
XX (REAS-) RES ASSOC BIOTECHNOLOGY.
XX
XX Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S,
XX Yamamoto J, Isono Y, Hio Y, Otsuka K, Nagai K, Irie R, Tamechika I,
XX Seki N, Yoshikawa T, Otsuka M, Nagahari K, Masuho Y;
XX WPI; 2003-723558/69.
XX P-PSDB; ADM03788.
XX
XX New polynucleotides and polypeptides are useful in gene therapy, for
XX developing a diagnostic marker or medicines for regulating their
XX expression and activity, or as a target of gene therapy.
XX
XX Claim 1; SEQ ID NO 30; 305pp; English.
XX
XX The invention relates to a novel human polynucleotide and the encoded
XX polypeptide. A polynucleotide of the invention may have a use in gene
XX therapy. An oligonucleotide of the invention ADM06202-ADM06773 is useful
XX as a primer for synthesizing the polynucleotide or as a probe for
XX detecting the polynucleotide. The polynucleotides ADM01316-ADM03758 are
XX useful in gene therapy, for developing a diagnostic marker or medicines
XX for regulating their expression and activity, or as a target of gene
XX therapy. The proteins ADM03759-ADM06201 encoded by the polynucleotides
XX are useful as pharmaceutical agents. The present sequence represents a
XX cDNA sequence of the invention.
XX
XX Sequence 2836 BP; 632 A; 722 C; 676 G; 806 T; 0 U; 0 Other;
Query Match 13.3%; Score 267.2; DB 11; Length 2836;
Best Local Similarity 78.8%; Pred. No. 1.5e-58;
Matches 350; Conservative 0; Mismatches 78; Indels 16; Gaps 2;
QY 292 GCGCGGTTGGTGGCTCATGCTCTAATCCAGCACTTTGGAGGCCAAGTGAGTGA 351
DB 2473 GGCTGGGCGACAGTGGCTCACACCTGTATATCCAGCACTCTAGGAGGCCAAGTGAGTGA 2414
QY 352 TCACCTTGAGGCGGGAGTTTGAGACGAGCTCGCCCAACATGTGAACCCCATCTGTACT 411
DB 2413 CCACCTTGAGTCAAGAGTTTCAGACCCAGCTCGGCGCAACCTCGTCTCTACT 2354
QY 412 AAAAATACAAAGTTAGTGGCGGTGGTGGTGTAGATGCCCTGTAGTCCCACTACTTTGGAG 471
DB 2353 GAAAATATAAAATTAGCCAGCATGTGGGCGACAGCTGTATCCAGCAGTGTGGAG 2294
QY 472 GCTGAGCATGAGATC-----GTTGAGCCCGCAGCTGGGCAATACAGCAA 517

CC comprising assessing the level of expression or activity of a PTPN11
CC protein in the test subject, and comparing it to the level of expression
CC or activity in a control subject, where an increased expression or basal
CC activity of the PTPN11 protein in the test subject compared to the
CC control is indicative of Noonan syndrome; (3) treating Noonan syndrome in
CC a patient, comprising administering an agent that modulates the
CC expression or activity of a PTPN11 protein in association with a carrier;
CC (4) an isolated PTPN11 variant comprising a mutation resulting in
CC increased level of PTPN11 activity; (5) an isolated cell comprising a
CC vector comprising a nucleic acid encoding the PTPN11 variant of (4),
CC operatively associated with an expression control sequence; (6) an
CC isolated nucleic acid encoding the PTPN11 variant of (4); and (7) an
CC isolated oligonucleotide which specifically hybridises to the nucleic
CC acid of (6). The methods and compositions of the present invention are
CC useful for diagnosing and treating a disorder associated with the
CC aberrant expression and/or activity of the PTPN11 gene, specifically
CC Noonan syndrome. The present sequence represents human PTPN11 genomic
CC DNA, which is given in the exemplification of the present invention.
XX
SQ Sequence 300000 BP; 84671 A; 64420 C; 64260 G; 85849 T; 0 U; 800 Other;
Query Match 13.2%; Score 265.4; DB 10; Length 300000;
Best Local Similarity 79.5%; Pred. No. 2.8e-57;
Matches 357; Conservative 0; Mismatches 76; Indels 16; Gaps 3;
QY 291 TGGCCGGGTGGTGGCTCATCCCTTAATCCAGCACCTTTGGGAGGCCAAGGTGAGTAG 350
Db 121055 TGGCCAAAGTGGGTGGCTTACACCTGTATCCAGCACCTTTGGGAGGCCAAGGTGGTGG 120996
QY 351 ATCACTTGAGCGCGGAGTTTGAGACCGACCTGGCCAAACATGGTGAACCCCATCTGTAC 410
Db 120995 ATCACTGAGGTTCAGAGTTTGAGACCGACCTGGCCAAACATGGTGAACCAATCTGTAC 120936
QY 411 TAAAAATACAAAGTTA--GCTGGGGGTGGTGGTAGATGCTGTAGTCCAGCTACTTGG 468
Db 120935 TAAAAATACAAAAATTTAGGGCTGGGCGGGTGGCTCACGCCGTGTAATCCAGCACCTTGG 120876
QY 469 GAGGCTGAGGCATGAGATC-----GCTTGAGCCCGAGCTGGGCAATACAGCA 516
Db 120875 GAGGCTGAGCGGGGGATGACAGAGTCAGGAGATCGAGACCATCTGGCTAACACAGTG 120816
QY 517 AGACCCCGCTCTCTACAAATAAATACAAAAATAGTTGGATGTGGTGTGATGCGCTGT 576
Db 120815 AAACCCCATCTCTACTA--AAATACAAAGATTAGCTGGGCGTGGTGGGCACTGT 120758
QY 577 AGTCTAGCTGTAGGAGGCTGAGATGAAGATTGCTTGGCTGGGAGGCTCAAGGCT 636
Db 120757 AGTCCCAAGCTGTCTGGGAGGCTGAGGAGGAGATGGCGTGAACCCAGGAGTGGAGCTT 120698
QY 637 GCAGTGAGCGGAGATGGCGCCACTGCACCTCCAGCCTGGGCAACAGAGTGAGACCCCTGTCT 696
Db 120697 GCAGTGAGCGGAGATGACACCATCTGCCTCCAGCCTGGGCAACAGAGTGAGACTGTCT 120638
QY 697 CAGAAAAAATAAAAAAAAAAAGGAGA 725
Db 120637 CAAAAAATAAAAAAAAAAATAAATA 120609
RESULT 10
AD014076/c
ID AD014076 standard; DNA; 300001 BP.
XX
AC AD014076;
XX
DT 12-AUG-2004 (first entry)
XX
DE Human protein tyrosine phosphatase 11 gene sequence SEQ ID NO:33.
XX
KW protein tyrosine phosphatase gene 11; PTPN11; enzyme;
KW protein tyrosine phosphatase gene 11 variant; PTPN11 variant;
KW haematologic disorder; mutation; increased PTPN11 activity; cytostatic;
KW neuroprotective; PTPN11 modulator; acute lymphoblastic leukaemia; ALL;
KW acute myeloid leukaemia; AML; juvenile myelomonocytic leukaemia; JMML;
KW

KW myelodysplastic syndrome; MDS; cancer; pre-cancerous condition;
KW lung cancer; colorectal cancer; pancreatic cancer; bladder cancer;
KW kidney cancer; thyroid cancer; melanoma; leukaemia; human; chromosome 12;
KW gene; ds.
XX Homo sapiens.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT exon 123211..123604
FT /*tag= a
FT /number= 1
FT CDS 123591..248877
FT /*tag= b
FT /product= "PTPN11"
FT /transl_except= (pos:246260..246262,aa:Arg)
FT /transl_except= (pos:246299..246301,aa:Pro)
FT intron 123605..136830
FT /*tag= c
FT /number= 1
FT exon 136831..136953
FT /*tag= d
FT /number= 2
FT intron 136954..194430
FT /*tag= e
FT /number= 2
FT exon 194431..194625
FT /*tag= f
FT /number= 3
FT intron 194626..197307
FT /*tag= g
FT /number= 3
FT exon 197308..197500
FT /*tag= h
FT /number= 4
FT intron 197501..198676
FT /*tag= i
FT /number= 4
FT exon 198677..198793
FT /*tag= j
FT /number= 5
FT intron 198794..200062
FT /*tag= k
FT /number= 5
FT exon 200063..200176
FT /*tag= l
FT /number= 6
FT intron 200177..217056
FT /*tag= m
FT /number= 6
FT exon 217057..217153
FT /*tag= n
FT /number= 7
FT intron 217154..221763
FT /*tag= o
FT /number= 7
FT exon 221764..221843
FT /*tag= p
FT /number= 8
FT intron 221844..221969
FT /*tag= q
FT /number= 8
FT exon 221970..222128
FT /*tag= r
FT /number= 9
FT intron 222129..226186
FT /*tag= s
FT /number= 9
FT exon 226187..226318
FT /*tag= t
FT /number= 10
FT intron 226319..230587
FT /*tag= u


```
XX 01-MAY-2003.
PD 22-OCT-2002; 2002WO-US033835.
PF 23-OCT-2001; 2001US-00004113.
XX (SAGR-) SAGRES DISCOVERY.
XX Engelhard EK, Morris DW;
XX WPI; 2003-421412/39.
XX New recombinant nucleic acid and its encoded protein, useful for
PT preparing a composition for diagnosing or treating carcinomas.
XX Claim 1; Page 103-107; 173pp; English.
XX The invention relates to novel sequences which are useful for preparing a
CC composition for diagnosing or treating carcinomas. These sequence are
CC also useful in gene therapy. The present sequence is human CCND3
CC carcinoma associated (CA) gene. This sequence is used in the invention
XX
XX Sequence 26874 BP; 6621 A; 6603 C; 6750 G; 6880 T; 0 U; 20 Other;
SQ
  Query Match      13.1%; Score 262.8; DB 8; Length 26874;
  Best Local Similarity 76.6%; Pred. No. 5.1e-57;
  Matches 340; Conservative 0; Mismatches 92; Indels 12; Gaps 1;
QY 292 GCGCGGGTGGTGCTCATGCTTAATCCAGCACTTTGGAGGCGCAAGTGAGTAGA 351
Db 5012 GGTGGGCGGTGGTCCACACCTTAATCCAGCACTTTGGAGGCGCGAGTGGCGGA 4953
QY 352 TCACCTGAGTGGAGTTCAGACCACTGGCCCAACATGGTGAACCCCATCTGTACT 411
Db 4952 TCACCTGAGTGGAGTTCAGACCACTGGCCCAACATGGTGAACCCCATCTGTACT 4893
QY 412 AAAAAATACAAAAGTTAGTGGCGGTGGTGGTAGATGCTGTAGTCCAGCTACTCTGGGAG 471
Db 4892 AAAAAATACAAAAGTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGGAG 4833
QY 472 GCTGAGGCGGAGTTCAGACCACTGGCCCAACATGGTGAACCCCATCTGTACT 411
Db 4832 GCTGAGGCGGAGTTCAGACCACTGGCCCAACATGGTGAACCCCATCTGTACT 4773
QY 520 CCCCCTCTCTACAAAATAAATAACAAAAAATTAGTTGGATGGTGGTGGTGGTGGTGGT 579
Db 4772 CCCCCTCTCTACAAAATAAATAACAAAAAATTAGTTGGATGGTGGTGGTGGTGGTGGT 4713
QY 580 CCTAGTCTGAGGAGGCTGAGATGGAAGGATTTGCTTGAAGCTGGGAGGTCAAGGCTGCA 639
Db 4712 CCAGCTACTCAGAGGCTGAGGAGGAGGAGTTCGCTTGAACCCACCAGGTAGAGTTGCA 4653
QY 640 GTGAGCGGAGTGGGCGCACTGCACTCCAGCTGGGCAACAGAGTGAGACCTGTCTGAG 699
Db 4652 GTGAGCTGAGATCAGCCCACTGCACTCCAGCTGGGCAACAGAGTGAGACCTGTCTGAG 4593
QY 700 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 723
Db 4592 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 4569
RESULT 12
IDA02467/c
ID ADA02467 standard; DNA; 26874 BP.
XX ADA02467;
AC
XX
XX
XX 06-NOV-2003 (first entry)
DE Human CCND3 carcinoma associated gene, SEQ ID NO:986.
XX Human; carcinoma associated; oncogene; carcinoma; cancer; breast;
```

```
KW prostate; lymphoma; leukaemia; cytostatic; gene therapy; drug screening;
KW gene; ds.
XX Homo sapiens.
OS
XX WO2003057146-A2.
PN
XX 17-JUL-2003.
PD
XX 26-DEC-2002; 2002WO-US041414.
PF
XX 26-DEC-2001; 2001US-00035832.
PR
XX (SAGR-) SAGRES DISCOVERY.
PA
XX Morris DW;
PI
XX WPI; 2003-587068/55.
DR
XX New recombinant nucleic acid encoding carcinoma associated protein,
PT useful for preparing compositions for treating carcinomas.
PT
XX Claim 1; SEQ ID NO 986; 245pp; English.
PS
XX The invention relates to recombinant carcinoma associated (CA) nucleic
CC acid sequences from mouse and human (ADA01482-ADA03094), and to
CC recombinant carcinoma associated proteins (CAP) encoded by them. The
CC invention also encompasses expression vectors and host cells comprising a
CC CA nucleic acid, a polypeptide (especially an antibody) that specifically
CC binds to the protein, and a biochip comprising CA nucleic acid or
CC fragments thereof. The sequences of the invention were identified using
CC oncogenic retroviruses, which insert into the genome of the host organism
CC at random. Many of these do not carry transduced host oncogenes or
CC pathogenic trans-acting viral genes, meaning that cancer incidence is a
CC direct consequence of the effects of proviral integration into host
CC carcinoma (especially breast cancer, prostate cancer, lymphoma or
CC protooncogenes). The CA nucleic acid sequences can be used to diagnose
CC leukaemia) or a propensity to carcinoma by determination of the sequence
CC of a CA gene, or by determination of CA gene expression in particular
CC tissues. CA nucleic acids, proteins and antibodies are also useful as
CC therapeutic agents and in screening and evaluating drug candidates. The
CC present sequence represents a specifically claimed human CA nucleic acid
CC sequence of the invention. Note: The complete sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 26874 BP; 5621 A; 6602 C; 6751 G; 6880 T; 0 U; 20 Other;
  Query Match      13.1%; Score 262.8; DB 9; Length 26874;
  Best Local Similarity 76.6%; Pred. No. 5.1e-57;
  Matches 340; Conservative 0; Mismatches 92; Indels 12; Gaps 1;
QY 292 GCGCGGGTGGTGCTCATGCTTAATCCAGCACTTTGGAGGCGCAAGTGAGTAGA 351
Db 5012 GGTGGGCGGTGGTCCACACCTTAATCCAGCACTTTGGAGGCGCGAGTGGCGGA 4953
QY 352 TCACCTGAGGCGGAGTTCAGACCACTGGCCCAACATGGTGAACCCCATCTGTACT 411
Db 4952 TCACCTGAGTGGAGTTCAGACCACTGGCCCAACATGGTGAACCCCATCTGTACT 4893
QY 412 AAAAAATACAAAAGTTAGTGGCGGTGGTGGTAGATGCTGTAGTCCAGCTACTCTGGGAG 471
Db 4892 AAAAAATACAAAAGTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGGAG 4833
QY 472 GCTGAGGCGGAGTTCAGACCACTGGCCCAACATGGTGAACCCCATCTGTACT 411
Db 4832 GCTGAGGCGGAGTTCAGACCACTGGCCCAACATGGTGAACCCCATCTGTACT 4773
QY 520 CCCCCTCTCTACAAAATAAATAACAAAAAATTAGTTGGATGGTGGTGGTGGTGGTGGT 579
Db 4772 CCCCCTCTCTACAAAATAAATAACAAAAAATTAGTTGGATGGTGGTGGTGGTGGTGGT 4713
```


CC	The invention relates to a novel human Ras-like protein comprising a	Query Match	13.0%;	Score 262;	DB 6;	Length 20951;
CC	fully defined sequence of 615 amino acids, given in the specification,	Best Local Similarity	70.0%;	Pred. No. 7.4e-57;		
CC	its allelic variant, orthologue or its fragment comprising at least 10	Matches 367;	Conservative	0;	Mismatches 155;	Indels 2;
CC	contiguous amino acids. The new human Ras-like proteins and the					Gaps 1;
CC	polypeptides encoding them are useful in the diagnosis, prevention, and					
CC	treatment of inflammation and disorders associated with cell					
CC	proliferation and apoptosis. They are also useful in treating AIDS and					
CC	other infectious or genetic immunodeficiencies, neurodegenerative					
<hr/>						
QY	292 GCGCGGTGTGGTCTCATGCTCTTAATCCAGCACTTTGGAGGCCAAGGTGAGTAGA	351				
DB	13494 GCGCAGGTGGGTGGCTCAGCCCTATAATCCAGAACTTCGGGAGGCGGCGGGA	13435				
QY	352 TCACCTGAGCGCGGAGTTTGAGACCAAGCTGGCCAAACATGCTGAACCCCATCTGTACT	411				
DB	13434 TCACCTGAGGTCAAGAGTTCCATACCAGCTTCGGCAACATGTGAACCCCATCTCTACT	13375				
QY	412 AAAAATACAAAAGTTAGCTGGCGGTGGGTAGATGCCTGTAGTCCCAGCTACTTTGGAG	471				
DB	13374 AAAAATACAAAATTAGCTGGGTGAGTGGCTCACGCTGTAATCCAGCACTTTGGAG	13315				
QY	472 GCTGAGGCATGAAATCGTTGAGCCAGCGCTGGGCAATACAGCAAGACCCCGTCTCTAC	531				
DB	13314 GCTGAGGCTAGTGGATCAAGAGTCTGGAGGTTCGAGACCAACCGTGAACCCCGTCTCTAC	13255				
QY	532 AAATAAAATACAAAAAATTAGTTGGATGTGGTGGTGCATGCGCTGTAGTCTAGCTGCTAG	591				
DB	13254 TA--AAATACAAAAAATTAGCCGGGCAAGTGTGGTGGTGTGTAGTCTCCAGTACTTCC	13197				
QY	592 GGAGGTGAGATGGAAAGGATTCGTTGAGCCTGGGAGGTCAAGGCTGCAGTCGAGCCGAGAT	651				
DB	13196 GGAGGCTGAGCGAGGAGATGGCATGAACCCGGGAAGCAGAGCTTGCAGTGAGCCGAGAT	13137				
QY	652 GCGGCCACTGCATCCAGCTGGGCAACAGATGAGACCCCTGTTTCAGAAAAAATAAAAA	711				
DB	13136 CCGCCACCCGCATCTCAGTCTCGGTTACAGATGAGACTCTGTCTCAAAAAAATAAAAAAT	13077				
QY	712 AAAAAAAGGAGGAGAGAGACTCAAGCAGCCGCTTCACAGGACTGCTGAGGCCCTGC	771				
DB	13076 ACAAATTAGCCAGGCATGTTGGCGCGCTGTAAATCCAGCTACTCGGAGGCTGAGA	13017				
QY	772 AGTGTCTGCAGCATGTGGCCCCAGCGCGGAGCTCTGTAAACC	815				
DB	13016 CAGGAGAATCACTTGAACCCGGGAGCGGAGGTTCGAGTGAACC	12973				

RESULT 15
AAK99739
ID AAK99739 standard; cDNA; 2568 BP.
XX
XX AAK99739;
XX
XX 11-JUL-2002 (first entry)
DT
XX
XX
DE cDNA encoding the angiotensin converting enzyme (ACE).
XX
XX Cardiovascular; vasotropic; angiotensin II; aortic valve disease;
KW angiotensin converting enzyme; aortic valve; aortic stenosis; ACE;
KW inhibiting agent; angiotensin II type 1 receptor antagonist;
KW aortic sclerosis; gene; ss.

OS	Homo sapiens.
XX	
PH	
Key	Location/Qualifiers
CDS	1155..1913
FT	/*tag= a
FT	/product= "Angiotensin converting enzyme"
FT	/transl_except= (pos:1170..1172, aa:Xaa)
FT	/transl_except= (pos:1308..1910, aa:Xaa)

FT		/note= "Xaa is unknown"
XX		
PN	WO200228421-A1.	
XX		
PD	11-APR-2002.	
XX		
PF	05-OCT-2001; 2001WO-US031605.	
XX		
PR	06-OCT-2000; 2000US-0238367P.	
XX	(UNIW) UNIV WASHINGTON.	
PA		
XX		
PI	O'brien KD, Otto CM, Probstfield JL;	
XX	WPI; 2002-372197/40.	
DR	P-PSDB; AA020950.	
XX		
PT	Decreasing the amount and/or biological	
PT	aortic valve comprises administering an	
PT	consisting of antagonists of angiotensin	
PT	angiotensin II type 1 receptor.	
XX		
PS	Disclosure; Page 30-32; 40pp; English.	
XX		
CC	The invention relates to a method for d	
CC	biological activity of angiotensin II i	
CC	This method comprises administering an	
CC	consisting of angiotensin converting en	
CC	angiotensin II type 1 receptor antagoni	
CC	prevent progression and/or complication	
CC	aortic stenosis or aortic sclerosis. The	
CC	represents the cDNA encoding the angiot	
CC	the invention	
XX		
SQ	Sequence 2568 BP; 558 A; 831 C; 712 G; ;	
<hr/>		
	Query Match	
	Best Local Similarity 13.0%; Score 261.	
	Matches 421; Conservative 0; Mismatch	
Qy	905 CCAGCTCTGAAATTCCTTGAGCTGCCCTTAT	
Db		
	604 CCAGCTCTGAAATTCCTTGAGCTGCCCTTAT	
Qy	965 AAGGCATTCAACACCCTACACAGATCTGACGG	
Db	664 AAGGCATTCAACACCCTACACAGATCTGACG	
Qy	1025 GACTGTATTGGCGATGGAGGGGTGGCGA	
Db	724 GACTGTATTGGCGATGGAGGGGTGGCGA	
Qy	1085 TACCGGAATACTGGAACCTCATCAACACG	
Db	784 TACCGGAATACTGGAACCTCATCAACACG	
Qy	1145 CCAACATCACTGGCAGCTGGGTGCCCTTCAT	
Db	829 -----	
Qy	1205 GCCTAGAAAAGGTAGATTCCTGGAGGAGG	
Db	829 -----	
Qy	1265 GTGTCCCTCTGTAGGCTATGTAGATGACG	
Db	829 -----GGCTATGTAGATGACG	
Qy	1325 CATCCCTGGAGCAAGACCTGAGCGGCTCTTT	
Db	875 CATCCCTGGAGCAAGACCTGAGCGGCTCTTT	
Qy	1385 TGCATGCTACGTGCGCGGGCCCTGCACGCG	

Db 935 TGCATGCTACGTGCGCGGCCCTGCACCGTCACTACGGGGCCGACATCAACCTGG 994
Qy 1445 AGGGGCCCCATTCTGCTCACCCTGCTGGGTAA 1475
Db 995 AGGGGCCCCATTCTGCTCACCCTGCTGGGGAA 1025

Search completed: November 22, 2004, 17:23:08
Job time : 901.868 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 22, 2004, 15:54:56 ; Search time 8217.69 Seconds
(without alignments)
12349.531 Million cell updates/sec

Title: US-10-017-724-1_COPY_3000_5784

Perfect score: 2785

Sequence: 1 gaggacttgacggatgg.....aattgtaaaagggtttct 2785

Scoring table: IDENTITY NUC

Gapop 10_0 , Gapext 1.0

Searched: 32822875 seqs, 18219865908 residues

Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : EST:*

1: gb_est1:*

2: gb_est2:*

3: gb_hc:*

4: gb_est3:*

5: gb_est4:*

6: gb_est5:*

7: gb_est6:*

8: gb_gss1:*

9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	831	29.8	939	5	BQ933301
2	791.8	28.4	891	5	BQ925776
3	749.8	26.9	763	7	BQ000113
4	749.2	26.9	937	5	BQ934054
5	715.4	25.7	757	6	CA449275
6	704.8	25.3	752	6	CD511356
7	704.2	25.3	746	5	BX114562
8	700.2	25.1	876	1	AU121101
9	694.8	24.9	721	4	BW178851
10	689.4	24.8	778	6	CA414667
11	685.8	24.6	707	7	CN274716
12	683.4	24.5	793	1	A1346330
13	682.8	24.5	748	4	BG680063
14	681.2	24.5	750	5	BUE18192
15	677.8	24.3	681	7	CN274719
16	675.8	24.3	756	5	BQ008487
17	663.2	23.8	706	2	BF109743
18	651.2	23.4	668	7	CN274711
19	649	23.3	693	3	BG673014
20	646.8	23.2	714	6	CA428685
21	645.2	23.2	660	1	A1419866
22	643	23.1	672	6	CD512572
23	638.8	22.9	660	6	CA431209
24	638.8	22.9	661	4	BM272400

C 25	635.8	22.8	673	5	BQ002730	UI-H-EI1-
C 26	632.6	22.7	706	2	AM190828	AW190828 x142C04.x
C 27	629.8	22.6	651	4	BG939270	BC939270 cn31e12.x
C 28	625.2	22.4	647	6	CA425254	CA425254 UI-H-DF0-
C 29	623.4	22.4	712	6	CA420699	CA420699 UI-H-FG0-
C 30	623.2	22.4	718	1	A1453687	A1453687 t339a07.x
C 31	622	22.3	648	6	CA421578	CA421578 UI-H-FG0-
C 32	620.2	22.3	558	1	AA149859	AA149859 z147e10.f
C 33	619.2	22.2	707	5	BUE27598	BUE27598 UI-H-FG0-
C 34	617.6	22.2	642	5	BQ019822	BQ019822 UI-H-ED0-
C 35	615.4	22.1	701	6	CA446830	CA446830 UI-H-ED1-
C 36	610	21.9	826	6	CD511191	CD511191 AGENCOURT
C 37	609.8	21.9	708	5	BM997135	BM997135 UI-H-ED0-
C 38	609	21.9	817	4	BI523839	BI523839 603051961
C 39	608.8	21.9	637	1	AA625208	AA625208 af70f01.f
C 40	608.8	21.9	560	5	BM988626	BM988626 UI-H-DH0-
C 41	607.8	21.8	3371	9	AY398790	AY398790 Homo sapi
C 42	606.8	21.8	610	7	CN481711	CN481711 hwl1c03.y
C 43	604.2	21.7	669	6	CA422281	CA422281 UI-H-FG0-
C 44	600.4	21.6	655	5	BM971657	BM971657 UI-CF-ECL
C 45	600	21.5	648	1	AI571506	AI571506 tr85a06.x

ALIGNMENTS

RESULT 1
BQ933301
LOCUS
DEFINITION
AGENCOURT 8744271 Lupski sciatic_nerve Homo sapiens cDNA clone
IMAGE:6205936 5', mRNA sequence.
ACCESSION
BQ933301
VERSION
BQ933301.1
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 939)
NIH-MGC http://mgc.ncbi.nlm.nih.gov/.
AUTHORS
National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE
Unpublished (1999)
JOURNAL
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Dr. James R. Lupeki
cDNA Library Preparation: Life Technologies, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
Cloning by: Agencourt Bioscience Corporation
Cloning Distribution: MSC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM13629 row: j column: 17
High quality sequence stop: 686.

FEATURES
Location/Qualifiers
1...939
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clones="IMAGE:6205936"
/sex="male"
/tissue_type="sciatic nerve"
/dev_stage="adult, 70 yr"
/lab_host="DH10B"
/clone_lib="Lupski sciatic nerve"
/note="Vector: pCMV-SPORT6 (Life Technologies); Site_1:
Not1; Site 2: SalI; cDNA made by oligo-dT priming.
Directionally cloned using the following adaptors:
5'-TCGACCTACGCTCG-3' and
5'-GACTAGTCTAGCTCGAGCGCCGCT(15)-3'. Size selected >
1 kb for average insert length 1.87 kb. This is a primary
library, non-amplified. Library constructed by Life
Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor
College of Medicine) and is available through Life

[illegible]

NOTE: this is a NUM_POC library.									
ORIGIN	Query Match	26.9%;	Score 749.8;	DB 7;	Length 763;				
	Best Local Similarity	99.5%;	Pred. No. 1.3e-169;						
	Matches 751;	Conservative 0;	Mismatches 4;	Indels 0;	Gaps 0;				
QY	66	AGATATTGATGATGCTGTCCTGAAAACAATGCCATCAGTGAGACAGACTTCAGGAACCTT	125						
Db	8	AGATATTGATGATGCTGTCCTGAAAACAATGCCATCAGTGAGACAGACTTCAGGAACCTT	67						
QY	126	CCAGATGGTCCCTTGGATCCCAAGGGACACCCAAATTGATCCCACTGGGTCACTTCG	185						
Db	68	CCAGATGGTCCCTTGGATCCCAAGGGACACCCAAATTGATCCCACTGGGTCACTTCG	127						
QY	186	CCATCAAGGCAAGGAGCTGGTTCAGACAGCAACTCGGACCCCGGCATCGCTGTAGGTTT	245						
Db	128	CCATCAAGGCAAGGAGCTGGTTCAGACAGCAACTCGGACCCCGGCATCGCTGTAGGTTT	187						
QY	246	TGACGAGTTTGGGTCTGTGGACTTCAGTGGCAATTTACGTAAACACTGACCGGACGA	305						
Db	188	TGACGAGTTTGGGTCTGTGGACTTCAGTGGCAATTTACGTAAACACTGACCGGACGA	247						
QY	306	CGACTATGCTGGCTTCGCTTTTGGTTTACCAGTCACGACGCGCTTCTATGTGATGTG	365						
Db	248	CGACTATGCGCGCTTCGCTTTTGGTTTACCAGTCACGACGCGCTTCTATGTGATGTG	307						
QY	366	GAAGCAGGTGACGCAGACCTACTGGGAGGACGAGCCACGCGGGCCTATGGGTACTCCGG	425						
Db	308	GAAGCAGGTGACGCAGACCTACTGGGAGGACGAGCCACGCGGGCCTATGGGTACTCCGG	367						
QY	426	CGTGTCCCTCAAGGTGGTGAACTCCAACGCGGGACGCGGAGACCTGAGGAAACGCGCT	485						
Db	368	CGTGTCCCTCAAGGTGGTGAACTCCAACGCGGGACGCGGAGACCTGAGGAAACGCGCT	427						
QY	486	GTGGCACACGGGAAACAGCCGGGCGAGTGGAACTTATGGCAGCGACCCACGAGAACAT	545						
Db	428	GTGGCACACGGGAAACAGCCGGGCGAGTGGAACTTATGGCAGCGACCCACGAGAACAT	487						
QY	546	TGGCTGGAGGACTACAGCGCCTATAGGTGGCACTTGACTCAGCGCCCAAGACCGGCTA	605						
Db	488	TGGCTGGAGGACTACAGCGCCTATAGGTGGCACTTGACTCAGCGCCCAAGACCTGCTA	547						
QY	606	CATCAGAGTCTTAGTGCAATGAAGGAAACAGGTCTATGGCAGACTCAGGACCTATCTATGA	665						
Db	548	CATCAGAGTCTTAGTGCAATGAAGGAAACAGGTCTATGGCAGACTCANGACCTATCTATGA	607						
QY	666	CCAAACCTACGCTGGCGGGCGCTGGGTCTATTGTCTTCTCTCAAGAAATGGTCTATTT	725						
Db	608	CCAAACCTACGCTGGCGGGCGGTGGGTCTATTGTCTTCTCTCAAGAAATGGTCTATTT	667						
QY	726	CTCAGACCTCAAGTACGAATGCAGAGATATTTAAACAAGATTGTGTCATTTCCGGCAAT	785						
Db	668	CTCAGACCTCAAGTACGAATGCAGAGATATTTAAACAAGATTGTGTCATTTCCGGCAAT	727						
QY	786	GCCCTGTGCATGCCATGTCCTTAGACACCTCAGT	820						

[illegible]

```

Db      728 GCNCTGTGTCATGCTGCTCCTAGACACCTCAGT 762

RESULT 4
BQ934054      937 bp      mRNA      linear      EST 21-AUG-2002
LOCUS      AGENCOURT 8817273 Lupski_sciatic_nerve Homo sapiens cDNA clone
DEFINITION      IMAGE:6202836 5', mRNA sequence.
ACCESSION      BQ934054
VERSION      BQ934054.1 GI:22349437
KEYWORDS      EST.
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 937)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Dr. James R. Lupski
cDNA Library Preparation: Life Technologies, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
Cloning Strategy: Agencourt Bioscience Corporation
Cloning Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Place: LLNL3621 row: i column: 13
High quality sequence stop: 613.
FEATURES             source
     1..937
     /organism="Homo sapiens"
     /mol_type="mRNA"
     /db_xref="taxon:9606"
     /clone="IMAGE:6202836"
     /sex="male"
     /tissue_type="sciatic nerve"
     /dev_stage="adult, 70 yr"
     /lab_host="DH10B"
     /clone_lib="Lupski_sciatic_nerve"
     /note="Vector: pCMV-SPORT6 (Life Technologies); Site 1:
     NotI; Site 2: SalI; cDNA made by oligo-dT priming.
     Directionally cloned using the following adaptors:
     5'-TCGACCCAGGCTCCG-3' and
     5'-GACTAGTCTAGATCGGAGCGGCCCT(15)-3'. Size selected >
     1 kb for average insert length 1.87 kb. This is a primary
     library, non-amplified. Library constructed by Life
     Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor
     College of Medicine) and is available through Life
     Technologies."
ORIGIN
Query Match      26.9%; Score 749.2; DB 5; Length 937;
Best Local Similarity 99.0%; Pred. No. 1.9e-169;
Matches 754; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY      1126 ATAATCTGTCATAGCCATTATGATAAATTAAGCATGAATAATTTGCTGAACTACTTTT 1185
DB      1 AATACTGTGCAAGCCATTATGATAAATTAAGCATGAATAATTTGCTGAACTACTTTT 60

QY      1186 GGTGCTTAAAGTGTGCTACTTCTTGAATTAGAGTTGCTCTCAATGACACACAAATCCC 1245
DB      61 GTGTCTTAAAGTGTGCTACTTCTTGAATTAGAGTTGCTCTCAATGACACACAAATCCC 120

QY      1246 GCTAATAAATTAATAACAGGGTCAATTCAAATTTGAAGTAATGTTTGTAGTAAGAGAG 1305
DB      121 GCTAATAAATTAATAACAGGGTCAATTCAAATTTGAAGTAATGTTTGTAGTAAGAGAG 180

QY      1306 ATTAGAAGACACAGGCATAGCAATGACATAGCTACCGATTAACTAATCGGAACATGT 1365
DB      181 ATTAGAAGACACAGGCATAGCAATGACATAGCTACCGATTAACTAATCGGAACATGT 240

```

```

QY      1366 AAAACAGTTACAAAAATAAAACGAACTCTCTCTCTGTTCTTACAAATGAAGCCCTCATGTGC 1425
DB      241 AAAACAGTTACAAAAATAAAACGAACTCTCTCTCTGTTCTTACAAATGAAGCCCTCATGTGC 300

QY      1426 AGTAGAGATGAGTTTTCATCAAAAGAACAAACATCCTTGCRAATGGTGTGACGGGTTC 1485
DB      301 AGTAGAGATGAGTTTTCATCAAAAGAACAAACATCCTTGCRAATGGTGTGACGGGTTC 360

QY      1486 AGATGTGAGTTTGGCAAAACCTCATTTAAAGTAAAGGTTAGCAGACGAAAGTGGGTGCT 1545
DB      361 AGATGTGAGTTTGGCAAAACCTCATTTAAAGTAAAGGTTAGCAGACGAAAGTGGGTGCT 420

QY      1546 TTAGCTGCTGCTTGTGCGGTTGTGCGTGGGAGGCTCTGCTGAGCTTCTTCCCA 1605
DB      421 TTAGCTGCTGCTTGTGCGGTTGTGCGTGGGAGGCTCTGCTGAGCTTCTTCCCA 480

QY      1606 GCTTTGCTGCTGAGAGACGAGACGAGACGACGCGGCGGAAAGCGCATCTAACGC 1665
DB      481 GCTTTGCTGCTGAGAGACGAGACGAGACGCGGCGGCGGAAAGCGCATCTAACGC 540

QY      1666 GTATCTAGGCTTTGTGTAACCTCGGACAAAGTTGCTTTTACCTGATTGTGATGATCATTTCA 1725
DB      541 GTATCTAGGCTTTGTGTAACCTCGGACAAAGTTGCTTTTACCTGATTGTGATGATCATTTCA 600

QY      1726 TTAGGTTCCAGTTATAAATTTTGTAAATTTTAAATTTTAAAGTACTATAGTAACTCACTC 1785
DB      601 TTAGGTTCCAGTTATAAATTTTGTAAATTTTAAATTTTAAAGTACTATAGTAACTCACTC 660

QY      1786 CATTTACCAGTAACCTATTATTAATATGCTAGTAACACATATGATGATTAATTTCTAGA 1845
DB      661 CATTTACCAGTAACCTATTATTAATATGCTAGTAACACATATGATGATTAATTTCTAGA 720

QY      1846 AACAAACATCTAATAAGTATATATCTCTGTGAAATATGAGG 1887
DB      721 AACAAACATCTAATAAGTATATATCTCTGTGAAATATGAGG 762

RESULT 5
CA449275/c
LOCUS      CA449275      757 bp      mRNA      linear      EST 08-NOV-2002
DEFINITION      UI-H-E11-ayt-a-09-0-UI-s1 NCI CGAP_E11 Homo sapiens cDNA clone
ACCESSION      CA449275
VERSION      CA449275.1 GI:24813695
KEYWORDS      EST.
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 757)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Dr. Jose Mercuende
cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Cloning Distribution: Clone distribution information can be obtained
from Dr. M. Bento Soares, bento-soares@uiowa.edu
The following repetitive elements were found in this cDNA
sequence: 1-27, >AT rich#Low_complexity
Seq primer: M13 FORWARD
POLYA=Yes.
FEATURES             Location/Qualifiers
     1..757
     /organism="Homo sapiens"
     /mol_type="mRNA"
     /db_xref="taxon:9606"
     /clone="UI-H-E11-ayt-a-09-0-UI"
     /tissue_type="Chondrosarcoma"

```



```

Db      658 TAGTCAATATTGGTACTGATGATCTATCTGCTGNATATGGAATCTTTTAATTCAAAGC 717
QY      2060 CTGAACCAATCAGCAATTAGTCTTGCCA 2089
Db      718 CTGAACA-CAATCAGCAATTAGTCTTGCCA 746

RESULT 8
LOCUS   AUI21101 HEMBB1 Homo sapiens cDNA clone HEMBB1002084 5', mRNA
DEFINITION
ACCESSION AUI21101
VERSION   AUI21101.1 GI:10936336
KEYWORDS EST.
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Ota,T., Nishikawa,T., Suzuki,Y., Ishii,S., Saito,K., Kawai,Y.,
Yamamoto,J., Wakamatsu,A., Nakamura,Y., Nagai,T., Sugano,S. and
Isogai,T.
HRI human cDNA project
Unpublished (2000)
Contact: Takao Isogai
Genomics Laboratory
Helix Research Institute
1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
Tel: 81-438-52-3975
Fax: 81-438-52-3986
Email: genomics@hri.co.jp
HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix
Research Institute; cDNA library construction: Department of
Virology, Institute of Medical Science, University of Tokyo, and
Helix Research Institute.
Location/Qualifiers
source
1..876
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="HEMBB1002084"
/tissue_type="whole embryo, mainly body"
/dev_stage="embryo, 10 weeks"
/clone_lib="HEMBB1"
/note="Vector: pWE18SFL3"

ORIGIN
Query Match 25.1%; Score 700.2; DB 1; Length 876;
Best Local Similarity 99.6%; Pred. No. 1.2e-157;
Matches 702; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      49 GACATGACACATCCAGATATTGATGATGTTGCTCTGAAACAATGCCATCAGTGAG 108
Db      1 GACATGACACATCCAGATATTGATGATGTTGCTCTGAAACAAGTGCCATCAGTGAG 60

QY      109 ACAGACTTCAGAACTTCAGATGTTCCCTTGGATCCCAAGGACCAACCAATTGAT 168
Db      61 ACAGACTTCAGAACTTCAGATGTTCCCTTGGATCCCAAGGACCAACCAATTGAT 120

QY      169 CCCAACTGGGTCAATCGCCATCAAGCAAGGAGTGGTTTCAGACAGCCAACTCGGACCCC 228
Db      121 CCCAACTGGGTCAATCGCCATCAAGCAAGGAGTGGTTTCAGACAGCCAACTCGGACCCC 180

QY      229 GGCATCGCTAGGTTTGAAGTTTGGTCTGFGACTTCAGTGGACATTTCTACGTA 288
Db      181 GGCATCGCTAGGTTTGAAGTTTGGTCTGFGACTTCAGTGGACATTTCTACGTA 240

QY      289 AACACTGACCGGACGAGCACTATGCTGCTTCGCTTTGGTTACCACTCAAGACGCCG 348
Db      241 AACACTGACCGGACGAGCACTATGCTGCTTCGCTTTGGTTACCACTCAAGACGCCG 300

QY      349 TTCTATGTGGTATGTGGAAGCAGGTGACGACACCTACTGGGAGACCAGCCACGCGG 408

```

```

Db      301 TTCTATGTGGTATGTGGAAGCAGGTGACGACACCTACTGGGAGACCAGCCACGCGG 360
QY      409 GCCTATGGCTACTCCGGCGGTGTCCTCAAGGTGGTGAACCTCCACCACGGGGAGCGGCGAG 468
Db      361 GCCTATGGCTACTCCGGCGGTGTCCTCAAGGTGGTGAACCTCCACCACGGGGAGCGGCGAG 420
QY      469 CACCTGAGGAAGCGCTGTGGCAGACACGGGGAACACGCCGGGGAGGTGCGAACCTTATGG 528
Db      421 CACCTGAGGAAGCGCTGTGGCAGACACGGGGAACACGCCGGGGAGGTGCGAACCTTATGG 480
QY      529 CACGACCCCGAGGAACATTGGCTGGAAGGACTACACGGCTATAGGTGGCACCTGACTCAC 568
Db      481 CACGACCCCGAGGAACATTGGCTGGAAGGACTACACGGCTATAGGTGGCACCTGACTCAC 540
QY      589 AGGCCCAAGACCGGCTACATCAGAGTCTTAGTGATGAAGGAAAAACAGGTCTATGGCAGAC 648
Db      541 AGGCCCAAGACTGGCTACATCAGAGTCTTAGTGATGAAGGAAAAACAGGTCTATGGCAGAC 600
QY      649 TCAGGACCTATCTATGACCAACCTACGCTGGCGGGCGGCTGGTCTATTGTCTTCTCT 708
Db      601 TCAGGACCTATCTATGACCAACCTACGCTGGCGGGCGGCTGGTCTATTGTCTTCTCT 660
QY      709 CAAGAAATGGTCTATTCTCAGACCTCAAGTACGAATGCAGAGAT 753
Db      661 CAAGAAATGGTCTATTCTCAGACCTCAAGTACGAATGCAGAGAT 705

RESULT 9
LOCUS   BM718851
DEFINITION
UI-E-EO1-ajc-j-23-0-UI.r1 UI-E-EO1 Homo sapiens cDNA clone
UI-E-EO1-ajc-j-23-0-UI 5', mRNA sequence.
ACCESSION BM718851
VERSION   BM718851.1 GI:19037242
KEYWORDS EST.
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Bonaldo,M.F., Lennon,G. and Soares,M.B.
Normalization and subtraction: two approaches to facilitate gene
discovery
Genome Res. 6 (9), 791-806 (1996)
97044477
8889548
Contact: Soares, MB
Coordinated Laboratory for Computational Genomics
University of Iowa
375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: bento-soares@uiowa.edu
Tissue Procurement: Dr. Gregg Hageman
cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
DNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Researchers may obtain clones from Research
Genetics (www.resgen.com).
Seq primer: M13 Reverse
Location/Qualifiers
1..721
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-E-EO1-ajc-j-23-0-UI"
/dev_stage="fetal eye"
/tissue_type="fetal eye"
/lab_host="DH10B (Life Technologies) (T1 phage resistant)"
/clone_lib="UI-E-EO1"
/note="Organ: eye; Vector: pTT3-Pac (Pharmacia) with a
modified polylinker; Site_1: EcoR I; Site_2: Not I;"

```

FEATURES
source

ORGANISM
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
 1 (bases 1 to 793)
AUTHORS
 NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
TITLE
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
JOURNAL
 Unpublished (1997)
COMMENT
 Contact: Robert Strausberg, Ph.D.
 Email: cgaps-r@mail.nih.gov
 Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
 Emmert-Buck, M.D., Ph.D.
 CDNA Library Preparation: M. Bento Soares, Ph.D.
 CDNA Library Arrayed by: Greg Lennon, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/LINL at:
www-bio.llnl.gov/bbrp/image/image.html
 Insert length: 1324 Std Error: 0.00
 Seq primer: -40UP from Gibco
 High quality sequence stop: 456.
 Location/Qualifiers
 1..793
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:1926422"
 /tissue_type="adenocarcinoma"
 /lab_host="DH10B"
 /clone_lib="NCI CGAP Co8"
 /notes="Organ: colon; Vector: pT73D-Pac (Pharmacia) with a
 modified polylinker; 1st strand cDNA was prepared from
 colon adenocarcinoma, and was then primed with a Not I -
 oligo(dT) primer. Double-stranded cDNA was ligated to Eco
 RI adaptors (Pharmacia), digested with Not I and cloned
 into the Not I and Eco RI sites of the modified pT73
 vector. Library is normalized. Library was constructed by
 Bento Soares and M. Fatima Bonaldo. "
ORIGIN

Query Match	24.58;	Score 683.4;	DB 1;	Length 793;
Best Local Similarity	95.8;	Pred. No. 1.4e-153;		
Matches 744;	Conservative	0;	Mismatches 28;	Indels 5; Gaps 4;

Qy	393	GGACGACCCACCGGGCGCTATGCTACTCGGGCGTCTCCCTCAAGTGGTGAACCTCCAC	452
Db	773	GGAGAACAGCCACCGCGCTAATGGTACTCGGNGTGT-CCTCAAGTGGTGA--CTCAC	717
Qy	453	CACGGGACCGGGGAGACCTGTAGGAACACCGCTGTGGCACAACGGGGAAACAGCCGGGGCA	512
Db	716	CACGGGACCGGGGAGACCTTAGGGAACGGGTGGTGGCACAACGGGGAAACACCCGGGGCA	657
Qy	513	GGTGGAACTTATGGCAGCACCCAGGAACATTGGCTGGGAAGACTACACGGGCTATAG	572
Db	656	GGTGGCAA--CTTATGGCAGCACCCAGGAACCATTTGGCTGGGAAGACTACACGGGCTATAG	598
Qy	573	GTGGCACTGACTCACAGGCCAAGACCGGCTACATCAGAGTCTTAGTGCATGAAGAAA	632
Db	597	GTGGCACTGACTCACAGGCCAAGACTGGCTACATCAGAGTCTTAGTGCATGAAGAAA	538
Qy	633	ACAGGTCATGGCAGACTCAGGACCTTATATGACAAACCTTACGCTGGCGGGCGGCTGGG	692
Db	537	ACAGGTCATGGCAGACTCAGGACCTTATGACAAACCTTACGCTGGCGGGCGGCTGGG	478
Qy	693	TCTATTGTCTTCTCAGAAATGGTCTATTCTCAGACCTCAAGTACGATCGACAGAGA	752
Db	477	TCTATTGTCTTCTCAGAAATGGTCTATTCTCAGACCTCAAGTACGATCGACAGAGA	418
Qy	753	TATTTAAACAAGATTGCTGCATTTCGGGCAATG--CCTGTGTCATGCCATGGTCCCTAGA	811
Db	417	TATTTAAACAAGATNFGTGCATTTCGGGCAATGCCCTGTGTCATGCCATGGTCCCTAGA	358
Qy	812	CACCTCAGTTTATTTGGTGGCTTCTCTCTTAGCAGCACCTCTCTTCCCTTTA	871

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 22, 2004, 15:08:21 ; Search time 8322.11 Seconds
(without alignments)
11427.340 Million cell updates/sec

Title: US-10-017-724-3_COPY_85400_87410

Perfect score: 2011

Sequence: 1 ccggccgcttgaaagagctt.....gaatgtggaagtagcggt 2011

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4526729 seqs, 23644849745 residues

Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : GenEmbl.*

1: gb_ba.*
2: gb_htg.*
3: gb_in.*
4: gb_om.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_scs.*
12: gb_sy.*
13: gb_un.*
14: gb_vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %		Length	DB	ID	Description
		Match	Length				
1	2011	100.0	98829	2	AC002345		AC002345 Homo sapi
c	2	2011	100.0	179509	9	AC113554	AC113554 Homo sapi
	3	1997.4	99.3	23424	9	AY436326	AY436326 Homo sapi
4	1667.6	82.9	24070	9	AF118569	AF118569 Homo sapi	
5	1233.6	61.3	2649	6	A28005	A28005 Human ACE g	
6	824.6	41.0	1856	6	AR000113	AR000113 Sequence	
7	824.6	41.0	1856	6	AR137381	AR137381 Sequence	
8	824.6	41.0	1856	6	BD213164	BD213164 Genes for	
9	824.6	41.0	1856	6	BD075256	BD075256 Methods f	
10	824.6	41.0	1856	6	BD075259	BD075259 Methods f	
11	824.6	41.0	1856	9	HSATICE	X62855 H.sapiens cat	
12	490.6	24.4	195863	2	AL596246	AL596246 Lemur cat	
13	299.8	14.9	141136	10	AL596246	AL596246 Mouse DNA	
c	14	299.8	14.9	217088	2	AC023518	AC023518 Mus muscu
c	15	277.2	13.8	213169	2	AC096435	AC096435 Rattus no
c	16	277.2	13.8	234451	2	AC096705	AC096705 Rattus no
17	273.8	13.6	188006	2	AC147422	AC147422 Oryctolag	
c	18	271.6	13.5	29488	9	AF503510	AF503510 Homo sapi
19	270.8	13.5	164824	2	AC023887	AC023887 Homo sapi	

ALIGNMENTS

RESULT 1
AC002345
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS

AC002345
Homo sapiens chromosome 17 clone CTB-20D5 map 17, 8 unordered
pieces.

AC002345.2 GI:13027555
HTG: HTGS_PHASE1; HTGS_FULTOP; HTGS_CANCELLED.

Homo sapiens

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 98829)

Birren,B., Linton,L., Nusbaum,C. and Lander,E.

Homo sapiens chromosome 17, clone CTB-20D5

Unpublished

2 (bases 1 to 98829)

Hawkins,T.L., Reeve,M.P., Christoffersen,A., Birren,B.W.,
Fasman,K.H., Lander,E.S., McKernan,K., Munro,C., Richardson,P.,
Barna,N., Chang,A., Cooke,P., Daly,M.J., Devon,K., Dewar,K.,
Forrest,C., Gage,D., Geraghty,K., Guitau,G., Hagos,B., Huang,J.,
Jacotot,L., Lane,M., Lee,K., MacKenzie,J., Marquis,N.,
McDermott,J., Molla,M., Moloney,N., Morrow,J., Nachman,A.,
Naylor,J., O'Connor,T., Olotu,A., Peterson,K., Rollins,G.,
Spencer,J., Stilwell,J., Stone,C., Strickland,C., Sydney,K.,
Tang,L., Traish,A., Wilmer,F., Zemtseva,I. and Zody,M.

Direct Submission

Submitted (17-JUL-1997) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA

3 (bases 1 to 98829)

Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
Anderson,S., Barna,N., Bastien,V., Boguslavsky,L., Boukhalter,B.,
Brown,A., Camarata,J., Campiano,A., Chang,J., Chazaro,B.,
Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A.,
Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S.,
Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,
Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
Hagos,B., Heaford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R.,
Jones,C., Kanat,A., Karatas,A., Kells,C., LaRocque,K.,
Lamarez,R., Landers,T., Lechoczy,J., Levine,R., Liu,G.,
MacLean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C.,
McCarthy,M., McEwan,P., McKernan,K., McPheeters,R., Meldrim,J.,

20 270.6 13.5 172474 2 AC118582
21 270.4 13.4 40946 9 AC002984
22 270.4 13.4 41369 9 CH19F24590
23 270.2 13.4 148624 2 AC084815
24 270.2 13.4 173026 2 AC103883
c 25 270.2 13.4 175051 9 AC010531
c 26 270.2 13.4 191635 2 AC137760
c 27 270.2 13.4 235286 9 AC136285
c 28 270.2 13.4 132933 9 AL137002
c 29 269.2 13.4 155645 2 AC021153
30 269.2 13.4 177962 9 AC008763
31 268.8 13.4 182547 9 AP001201
32 268.8 13.4 208726 9 AP006287
c 33 268.4 13.3 122302 9 AC003982
c 34 267.8 13.3 193609 9 AC024580
c 35 267.4 13.3 196533 9 CNS000YVI
c 36 267.2 13.3 2836 6 AX832906
c 37 267.2 13.3 2836 9 AK094129
c 38 267.2 13.3 126414 9 AC108084
c 39 267 13.3 39351 9 AC020945
c 40 267 13.3 217280 2 AC019337
c 41 266.8 13.3 163038 2 AC135542
c 42 266.6 13.3 145009 9 AC106742
c 43 266.6 13.3 152464 2 AC027421
c 44 266.6 13.3 157243 9 AL353622
c 45 266.6 13.3 157986 2 AC025860

AC118582 Oryctolag
AC002984 Human DNA
AD001527 Homo sapi
AC084815 Homo sapi
AC103883 Homo sapi
AC010531 Homo sapi
AC137760 Homo sapi
AC136285 Homo sapi
AL137002 Human DNA
AC021153 Homo sapi
AC008763 Homo sapi
AP001201 Homo sapi
AP006287 Homo sapi
AC003982 Homo sapi
AC024580 Homo sapi
AL096870 Human chr
AX832906 Sequence
AK094129 Homo sapi
AC108084 Homo sapi
AC020945 Homo sapi
AC019337 Homo sapi
AC135542 Pan trogl
AC106742 Homo sapi
AC027421 Homo sapi
AL353622 Human DNA
AC025860 Homo sapi

Meneus, L., Mihova, T., Mlenka, V., Murphy, T., Naylor, J., Nguyen, C., Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neil, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Raymond, C., Retta, R., Kieback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schnupack, R., Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Straus, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J., Toham, K., Traversman, A., Travis, N., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G., Zainoun, J., Zembek, D., Zimmer, A. and Zody, M.

Direct Submission

Submitted (27-MAR-2003) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
On Feb 21, 2001 this sequence version replaced gi:2262095.

All repeats were identified using RepeatMasker:

Smit, A.F.A. & Green, P. (1996-1997)

<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WtBR

Web site: <http://www-seq.wi.mit.edu>

Contact: sequence_submissions@genome.wi.mit.edu

----- Project Information

Center project name: L115

Center clone name: 20_D_5

* NOTE: This is a 'working draft' sequence. It currently consists of 8 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

* 1 30263: contig of 30263 bp in length

* 30264 30363: gap of 100 bp

* 30364 36119: contig of 5756 bp in length

* 36120 36219: gap of 100 bp

* 36220 40234: contig of 4015 bp in length

* 40235 40334: gap of 100 bp

* 40335 51410: contig of 11076 bp in length

* 51411 51510: gap of 100 bp

* 51511 56945: contig of 5435 bp in length

* 56946 57045: gap of 100 bp

* 57046 68655: contig of 9820 bp in length

* 68656 68965: gap of 100 bp

* 68966 70772: contig of 3807 bp in length

* 70773 70872: gap of 100 bp

* 70873 98829: contig of 27957 bp in length.

Location/Qualifiers

1..98829

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

/chromosome="17"

/map="17"

/clone="CTB-20D5"

/clone_lib="CITB Human BAC"

FEATURES

source

ORIGIN

Query Match 100.0%; Score 2011; DB 2; Length 98829;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2011; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCGGCCCGTTTGAAGAGCTACCCCGGACACAGACCCGACACAGATACCTCCAGC 60
Db CCGGCCCGTTTGAAGAGCTACCCCGGACACAGACCCGACACAGATACCTCCAGC 85459

QY 61 TCCCTCTCAACCCACCCCTTTCAGGGTTGGAGAACTTGGAGCATAACTTGTTCATGA 120
Db TCCCTCTCAACCCACCCCTTTCAGGGTTGGAGAACTTGGAGCATAACTTGTTCATGA 85519

QY 121 GGAATCTCCACCCAGAAATGGGTCTTTCTGGCCCCCAGCCAGCTCCACATTAGAACA 180

Db	85520	GGAATCTCCACCCAGAAATGGGTCTTTCTGGCCCCCAGCCAGCTCCACATTAGAACA	85579
QY	181	TGACAAATAGAGGGGAAATGGAAATAAAGAGAGAAACGGTTTCCAGGACAGGGTT	240
Db	85580	TGACAAATAGAGGGGAAATGGAAATAAAGAGAGAAACGGTTTCCAGGACAGGGTT	85639
QY	241	TGGCTACAAAGTTGTGGATGGGTACCCATGCGCAAGTGTGAGGGAGGCTGGCCGGTG	300
Db	85640	TGGCTACAAAGTTGTGGATGGGTACCCATGCGCAAGTGTGAGGGAGGCTGGCCGGTG	85699
QY	301	TGGTGGCTCATGCTCTTAATCCAGCACTTTGGGAGGCCAGGTGAGTAGATCACTTGA	360
Db	85700	TGGTGGCTCATGCTCTTAATCCAGCACTTTGGGAGGCCAGGTGAGTAGATCACTTGA	85759
QY	361	GCGGGAGTTTGACACCAGCTGGCCAACTGGTGAACCCCACTCTGTACTTAAATAACA	420
Db	85760	GCGGGAGTTTGACACCAGCTGGCCAACTGGTGAACCCCACTCTGTACTTAAATAACA	85819
QY	421	AAAGTTAGCTGGCGTGGTGTAGATGCTGTAGTCCAGCTACTTGGAGAGCTGAGGCA	480
Db	85820	AAAGTTAGCTGGCGTGGTGTAGATGCTGTAGTCCAGCTACTTGGAGAGCTGAGGCA	85879
QY	481	TGAGAAATCGCTTGAGCCAGCTGGGCAATACAGCAAGACCCCGCTCTCAATAAATAA	540
Db	85880	TGAGAAATCGCTTGAGCCAGCTGGGCAATACAGCAAGACCCCGCTCTCTCAATAAATAA	85939
QY	541	ACAAAAATAGTTGGATGTGGTGTGATGCTCTGTAGTCTGTAGTCTGGAGGCTGA	600
Db	85940	ACAAAAATAGTTGGATGTGGTGTGATGCTCTGTAGTCTGTAGTCTGGAGGCTGA	85999
QY	601	GATGAAAGGATGTGTGAGCTGGAGGTCAAGGCTGAGTGGAGGAGTGGCGCACT	660
Db	86000	GATGAAAGGATGTGTGAGCTGGAGGTCAAGGCTGAGTGGAGGAGTGGCGCACT	86059
QY	661	GCATCTCAGCTGGGCAACAGAGTGAAGCCCTGTCTCAGAAAAAATAAAAAA	720
Db	86060	GCATCTCAGCTGGGCAACAGAGTGAAGCCCTGTCTCAGAAAAAATAAAAAA	86119
QY	721	GGAGAGGAGAGAGACTCAAGCAGCCCTCACAGGACTGTCTGAGGCCCTCAGAGTGTCTG	780
Db	86120	GGAGAGGAGAGAGACTCAAGCAGCCCTCACAGGACTGTCTGAGGCCCTCAGAGTGTCTG	86179
QY	781	CAGCATGTGGCCCGGAGGAGTCTGTGAAGCACTGTCTGGAGAGCACTCCCATCT	840
Db	86180	CAGCATGTGGCCCGGAGGAGTCTGTGAAGCACTGTCTGGAGAGCACTCCCATCT	86239
QY	841	TTCTCCAGATTTCTTAGACCTGCTGCTATACAGTCACTTTTATGTGTTTTCGCCAATTT	900
Db	86240	TTCTCCAGATTTCTTAGACCTGCTGCTATACAGTCACTTTTATGTGTTTTCGCCAATTT	86299
QY	901	TATTCAGCTCTGAAATTTCTCTGAGCTCCCTTTACAAGCAGAGGTGAGTAAAGGCTGA	960
Db	86300	TATTCAGCTCTGAAATTTCTCTGAGCTCCCTTTACAAGCAGAGGTGAGTAAAGGCTGA	86359
QY	961	GCTCAAGCATTTCAAAACCCCTACAGATCTGACGAATGTGTGGCCAGTCCCGGAAATA	1020
Db	86360	GCTCAAGCATTTCAAAACCCCTACAGATCTGACGAATGTGTGGCCAGTCCCGGAAATA	86419
QY	1021	TGAACACCTTTATGGCATGGGAGGCTGGGAGACAGAGGGGGAGAGCCATCTCCA	1080
Db	86420	TGAACACCTTTATGGCATGGGAGGCTGGGAGACAGAGGGGGAGAGCCATCTCTCCA	86479
QY	1081	GTTTTACCCGAAATACGTGGAACTCATCAACAGGCTGCCCGGCTCAATGTGTGAGTCCCT	1140
Db	86480	GTTTTACCCGAAATACGTGGAACTCATCAACAGGCTGCCCGGCTCAATGTGTGAGTCCCT	86539
QY	1141	GCTGCAACATCACTGGCACTTGGGTCCCTTCATTTCTCTCAAGAGGTGTGTGAACC	1200
Db	86540	GCTGCAACATCACTGGCACTTGGGTCCCTTCATTTCTCTCAAGAGGTGTGTGAACC	86599
QY	1201	CCAGCCCTAGGAAAGGTAGATCCCTGGAGGAGGAGGTAAATGTGTGTGGAGAGCCT	1260

```
Db      86600  CCAAGCCTAGAAAAGGTAGATCCCTGGAGGAGCAGGTAATGTGTTGGGAGAGCCT 86659
QY      1261  GGCCTGTGTCCTCTGAGCTATGTAGATGACAGGGAGCTCGTGGAGCTTATGTACGAG 1320
Db      86660  GGCCTGTGTCCTCTGAGCTATGTAGATGACAGGGAGCTCGTGGAGCTTATGTACGAG 86719
QY      1321  ACACCATCCCTGGAGCAAGACCTTGGAGCGGCTCTTCCAGGAGCTGCAGGCCTCTACCTC 1380
Db      86720  ACACCATCCCTGGAGCAAGACCTTGGAGCGGCTCTTCCAGGAGCTGCAGGCCTCTACCTC 86779
QY      1381  AACCTGCATGCTACGTCGCGCGGGCCCTGCACCGTCACTACGCGGCCAGCAGCATCAAC 1440
Db      86780  AACCTGCATGCTACGTCGCGCGGGCCCTGCACCGTCACTACGCGGCCAGCAGCATCAAC 86839
QY      1441  CTGAGGGGGCCCATTCCTGCTCACTGCTGGGTAAAGGGACACATGTCGGGCTTGAGGAG 1500
Db      86840  CTGAGGGGGCCCATTCCTGCTCACTGCTGGGTAAAGGGACACATGTCGGGCTTGAGGAG 86899
QY      1501  GTAAGACGACACACAGTGTGAGTGAGGGTTGGGACAGGGCTGACTAGAGGGTAGGGAGC 1560
Db      86900  GTAAGACGACACACAGTGTGAGTGAGGGTTGGGACAGGGCTGACTAGAGGGTAGGGAGC 86959
QY      1561  AGGCTGGGACTGAGAGACTCAGCCCTGTGGGGGATGGTTCGCCAGGCTGAGGGGGGT 1620
Db      86960  AGGCTGGGACTGAGAGACTCAGCCCTGTGGGGGATGGTTCGCCAGGCTGAGGGGGGT 87019
QY      1621  GGGCGCTGGAGTGGGAGCCCCCCTGTCATCTGTCGTCGTCGTCGTCGTCGTCGTCGTC 1680
Db      87020  GGGCGCTGGAGTGGGAGCCCCCCTGTCATCTGTCGTCGTCGTCGTCGTCGTCGTCGTC 87079
QY      1681  GTCCGCAAGTCAACATGAGTGGGGAAGGTAATATCTTGTCCAGGAGACACGCGC 1740
Db      87080  GTCCGCAAGTCAACATGAGTGGGGAAGGTAATATCTTGTCCAGGAGACACGCGC 87139
QY      1741  ACCATCAACATGTTGTGATCTTGAAGGGGAGAGAGGCTGTGAGTGGAGCTGGG 1800
Db      87140  ACCATCAACATGTTGTGATCTTGAAGGGGAGAGAGGCTGTGAGTGGAGCTGGG 87199
QY      1801  GAGGCTTTCCCAAGAGTGGCTGTGAGCAGGAGGCTCGGAAGATGACAGGGTTGACAGA 1860
Db      87200  GAGGCTTTCCCAAGAGTGGCTGTGAGCAGGAGGCTCGGAAGATGACAGGGTTGACAGA 87259
QY      1861  TGGAGTGGGGATGAGAGGACAGCGAGTGTTCAGGCCAAGGGAATGGAAACAAAGA 1920
Db      87260  TGGAGTGGGGATGAGAGGACAGCGAGTGTTCAGGCCAAGGGAATGGAAACAAAGA 87319
QY      1921  AGAAGCTGAGATGTAATCTACTTCAACCTGGAGCCCTCTTTGCAAGGGCTGCAATC 1980
Db      87320  AGAAGCTGAGATGTAATCTACTTCAACCTGGAGCCCTCTTTGCAAGGGCTGCAATC 87379
QY      1981  TCAGATGCCCTGAATGTGGAAGTAGGCGGT 2011
Db      87380  TCAGATGCCCTGAATGTGGAAGTAGGCGGT 87410

RESULT 2
AC113554/c
LOCUS      Homo sapiens chromosome 17, clone CTD-2501B8, complete sequence.
DEFINITION AC113554
ACCESSION  AC113554
VERSION    AC113554.9
KEYWORDS   HTG.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
            Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Eukaryota; Euthera; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 179509)
AUTHORS    Birren,B., Nusbaum,C. and Lander,E.
TITLE      Homo sapiens chromosome 17, clone CTD-2501B8
JOURNAL    Unpublished
REFERENCE  2 (bases 1 to 179509)
AUTHORS    Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
            Anderson,S., Barna,N., Bastien,V., Boguslavkiy,L., Boukhgalter,B.,
            Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B.,
            Chopel,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A.,
            Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S.,
            Ferreira,P., FitzHugh,W., Gage,D., Gallagan,J., Gardyna,S.,
            Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
            Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
            Kamat,A., Karatas,A., Kells,C., LaRoque,K., Lamazares,R.,
            Landers,T., Lebecky,J., Levine,R., Liu,G., Maclean,C., McCarthy,M.,
            Macdonald,P., Major,J., Marquis,N., Matthews,C., McCarthy,M.,
            McEwan,P., McKernan,K., Meldrim,J., Meneus,L., Mihova,T.,
            Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C.,
            Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J.,
            Peterson,K., Phunkhang,P., Pierre,N., Pollara,V., Raymond,C.,
            Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J.,
            Rosetty,M., Roy,A., Santos,R., Schauer,S., Schuback,R., Seaman,S.,
            Severly,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
            Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
            Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
            Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
            Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
            Direct Submission
            Submitted (01-MAR-2002) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
            3 (bases 1 to 179509)
            Birren,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S.,
            Barna,N., Bastien,V., Bloom,T., Boguslavkiy,L., Boukhgalter,B.,
            Camarata,J., Chang,J., Chazaro,B., Chopel,Y., Collymore,A.,
            Cooke,P., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
            Faro,S., Ferreira,P., FitzGerald,M., Gage,D., Gallagan,J.,
            Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hagos,B.,
            Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A.,
            Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K.,
            Liu,G., Maclean,C., Macdonald,P., Major,J., Matthews,C.,
            McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V.,
            Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H.,
            O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K.,
            Phunkhang,P., Pierre,N., Raymond,C., Retta,R., Rise,C., Rogov,P.,
            Roman,J., Roy,A., Schauer,S., Schuback,R., Seaman,S., Severly,P.,
            Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Talamas,J.,
            Tesfaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H.,
            Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J.,
            Zembek,L., Zimmer,A. and Zody,M.
            Direct Submission
            Submitted (16-AUG-2002) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
            4 (bases 1 to 179509)
            Birren,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S.,
            Barna,N., Bastien,V., Bloom,T., Boguslavkiy,L., Boukhgalter,B.,
            Camarata,J., Chang,J., Chazaro,B., Chopel,Y., Collymore,A.,
            Cooke,P., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
            Faro,S., Ferreira,P., FitzGerald,M., Gage,D., Gallagan,J.,
            Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hagos,B.,
            Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
            Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R.,
            Lindblad-Toh,K., Liu,G., Maclean,C., Macdonald,P., Major,J.,
            Matthews,C., McCarthy,M., Meldrim,J., Meneus,L., Mihova,T.,
            Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C.,
            Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J.,
            Peterson,K., Phunkhang,P., Pierre,N., Raymond,C., Retta,R.,
            Rise,C., Rogov,P., Roman,J., Roy,A., Schauer,S., Schuback,R.,
            Seaman,S., Severly,P., Smith,C., Spencer,B., Stange-Thomann,N.,
            Stojanovic,N., Talamas,J., Tesfaye,S., Theodore,J., Topham,K.,
            Travers,M., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X.,
            Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
            Direct Submission
            Submitted (26-OCT-2002) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
            On Oct 26, 2002 this sequence version replaced gi:22267826.
            All repeats were identified using RepeatMasker:
            Smit,A.F.A. & Green, P. (1996-1997)
            http://ftp.genome.washington.edu/RM/RepeatMasker.html
            ----- Genome Center
            Center: Whitehead Institute/ MIT Center for Genome Research
```

Center code: WBIR
Web site: <http://www-seq.wi.mit.edu>
Contact: sequence.submissions@genome.wi.mit.edu
-----Project Information
Center project name: L26087
Center clone name: 2501 B 8

FEATURES	source	Location/Qualifiers
repeat_region	1..179509	/organism="Homo sapiens"
repeat_region	/mol_type="genomic DNA"	
repeat_region	/db_xref="taxon:9606"	
repeat_region	/chromosome="17"	
repeat_region	/map="17"	
repeat_region	/clone="CTD-2501B8"	
repeat_region	/clone_lib="CITD2 Human BAC"	
repeat_region	116..324	
repeat_region	/rpt_family="MIR"	
repeat_region	complement(716..1014)	
repeat_region	/rpt_family="AluSx"	
repeat_region	complement(1392..1690)	
repeat_region	/rpt_family="AluSx"	
repeat_region	complement(1715..2012)	
repeat_region	/rpt_family="AluJb"	
repeat_region	complement(2023..2322)	
repeat_region	/rpt_family="AluSx"	
repeat_region	complement(2483..2791)	
repeat_region	/rpt_family="AluY"	
repeat_region	2921..3229	
repeat_region	/rpt_family="AluSx"	
repeat_region	complement(3346..3454)	
repeat_region	/rpt_family="MIR"	
repeat_region	complement(3455..3774)	
repeat_region	/rpt_family="AluSx"	
repeat_region	complement(3775..3789)	
repeat_region	/rpt_family="MIR"	
repeat_region	complement(5114..5177)	
repeat_region	/rpt_family="MIR"	
repeat_region	5577..5875	
repeat_region	/rpt_family="AluSg"	
repeat_region	5876..5896	
repeat_region	/rpt_family="GA)n"	
repeat_region	5928..6072	
repeat_region	/rpt_family="MIR"	
repeat_region	6083..6105	
repeat_region	/rpt_family="AT-rich"	
repeat_region	6118..6267	
repeat_region	/rpt_family="AluSq/x"	
repeat_region	complement(6426..7050)	
repeat_region	/rpt_family="L1MB8"	
repeat_region	complement(7056..7198)	
repeat_region	/rpt_family="FRAM"	
repeat_region	complement(7199..7376)	
repeat_region	/rpt_family="L1MB8"	
repeat_region	7377..7675	
repeat_region	/rpt_family="AluY"	
repeat_region	complement(7768..8412)	
repeat_region	/rpt_family="LTR65"	
repeat_region	8644..8799	
repeat_region	/rpt_family="MLT1J2"	
repeat_region	complement(8828..9134)	
repeat_region	/rpt_family="AluSx"	
repeat_region	9139..9424	
repeat_region	/rpt_family="MLT1I"	
repeat_region	9362..9656	
repeat_region	/rpt_family="AluSp"	
repeat_region	10157..10472	
repeat_region	/rpt_family="L2"	
repeat_region	10593..10745	
repeat_region	/rpt_family="L1PB2"	
repeat_region	10752..11059	
repeat_region	/rpt_family="AluSg"	
repeat_region	11091..11115	

Db 117663 AAAGTTAGCTGGCGCTGGTGTAGATCCCTGTAGTCCCACTACTTTGGGAGGCTGAGGCA 117604
QY 481 TGAGAAATCGCTTCAGAGCCAGCCCTGGCAATACAGCAAGACCCCGTCTCTACAAATAAAT 540
Db 117603 TGAGAAATCGCTTCAGAGCCAGCCCTGGCAATACAGCAAGACCCCGTCTCTACAAATAAAT 117544
QY 541 ACAAATAATAGTTGGATGTGGTGTGGTGCATGCTGTAGTCCCTAGCTAGGAGGCTGA 600
Db 117543 ACAAATAATAGTTGGATGTGGTGTGGTGCATGCTGTAGTCCCTAGCTAGGAGGCTGA 117484
QY 601 GATGGAAGATTCCTTGAGCTGGAGGCTCAAGGCTGCAAGGCTGAGCCGAGATGCGCCACT 660
Db 117483 GATGGAAGATTCCTTGAGCTGGAGGCTCAAGGCTGCAAGGCTGAGCCGAGATGCGCCACT 117424
QY 661 GCACCTCAGGCTGGGCAACAGAGTGAGAGCCCTCTCTCAGAAAAAATAAAAAAATAA 720
Db 117423 GCACCTCAGGCTGGGCAACAGAGTGAGAGCCCTCTCTCAGAAAAAATAAAAAAATAA 117364
QY 721 GGAGAGGAGAGACTCAAGCAGCCGCTTCAAGGACTGCTGAGGCGCTGCGAGGTGCTG 780
Db 117363 GGAGAGGAGAGACTCAAGCAGCCGCTTCAAGGACTGCTGAGGCGCTGCGAGGTGCTG 117304
QY 781 CAGCATGTGGCCCCAGGCGGGGACTCTGTAAGCCACTGCTGAGAGGCACTCCCATCT 840
Db 117303 CAGCATGTGGCCCCAGGCGGGGACTCTGTAAGCCACTGCTGAGAGGCACTCCCATCT 117244
QY 841 TTCTCCCATTTCTTAGACTGCTGCTATACAGTCACTTTTATGTGTTTCCCAATTT 900
Db 117243 TTCTCCCATTTCTTAGACTGCTGCTATACAGTCACTTTTATGTGTTTCCCAATTT 117184
QY 901 TATTCAGCTCTGAATTTCTTAGCTCCCTTCAAGCAGAGGTGAGCTTAAGGGCTGA 960
Db 117183 TATTCAGCTCTGAATTTCTTAGCTCCCTTCAAGCAGAGGTGAGCTTAAGGGCTGA 117124
QY 961 GCTCAGGCAATCAACCCCTACAGATCTGAGAAATGTAGGCACTGTCGCGAAATA 1020
Db 117123 GCTCAGGCAATCAACCCCTACAGATCTGAGAAATGTAGGCACTGTCGCGAAATA 117064
QY 1021 TGAAGACCTGTTATGGCATGGAGGCTGGGAGACAGGCGGGGAGAGCCATCTCCA 1080
Db 117063 TGAAGACCTGTTATGGCATGGAGGCTGGGAGACAGGCGGGGAGAGCCATCTCCA 117004
QY 1081 GTTTTACCCGAAATACGTGGAACTCATCAACAGGCTGCCCGCTCAATGGTGGTCCCT 1140
Db 117003 GTTTTACCCGAAATACGTGGAACTCATCAACAGGCTGCCCGCTCAATGGTGGTCCCT 116944
QY 1141 GCTGCCAATCACTGGCACTGGTCCCTTCAATTCCTCAAGAGGTGCTGTGAAC 1200
Db 116943 GCTGCCAATCACTGGCACTGGTCCCTTCAATTCCTCAAGAGGTGCTGTGAAC 116884
QY 1201 CCAAGCCTAGGAAAGGTAGATCCCTGGAGGAGGCAAGTAAATGTTGGGAGAGCCT 1260
Db 116883 CCAAGCCTAGGAAAGGTAGATCCCTGGAGGAGGCAAGTAAATGTTGGGAGAGCCT 116824
QY 1261 GCGTGTGCTCCCTCTGTAGGCTATGTAGATGAGGAGGCACTGCTGGAGGCTATGTACGAG 1320
Db 116823 GCGTGTGCTCCCTCTGTAGGCTATGTAGATGAGGAGGCACTGCTGGAGGCTATGTACGAG 116764
QY 1321 ACACCATCCTGGAGCAACCTGGAGGCTCTTCCAGAGGCTGAGGCACTTACCTC 1380
Db 116763 ACACCATCCTGGAGCAACCTGGAGGCTCTTCCAGAGGCTGAGGCACTTACCTC 116704
QY 1381 AACCTCATGCTTACCTGCGCGGCGCTGCAACCTCACTACGGGCGCCAGCACATCAAC 1440
Db 116703 AACCTCATGCTTACCTGCGCGGCGCTGCAACCTCACTACGGGCGCCAGCACATCAAC 116644
QY 1441 CTGGAGGGGCCCATTCCTGCTCACTGCTGGGTAAAGGCAATGTCGGGCTTGGAGAGG 1500
Db 116643 CTGGAGGGGCCCATTCCTGCTCACTGCTGGGTAAAGGCAATGTCGGGCTTGGAGAGG 116584
QY 1501 GTAAGACGGAACACAGTGTAGGTGGAGGCTGGGACAGGCTGACTAGAGGGTAGGGAGC 1560
Db 116583 GTAAGACGGAACACAGTGTAGGTGGAGGCTGGGACAGGCTGACTAGAGGGTAGGGAGC 116524

QY 1561 AGGCTGGGACTGAGAGACTCCAGCCCTGTGGGGATGCTGCCAGGCTGGAGGGGGT 1620
Db 116523 AGGCTGGGACTGAGAGACTCCAGCCCTGTGGGGATGCTGCCAGGCTGGAGGGGGT 116464
QY 1621 GGGCGCTGGAGTGGGGAGCCCCCCTTGCATCTGGTGCCACATTCATCTGCAGATCTAT 1680
Db 116463 GGGCGCTGGAGTGGGGAGCCCCCCTTGCATCTGGTGCCACATTCATCTGCAGATCTAT 116404
QY 1681 GTGCGGCAAGTCAACATGATGGGGAGAAAGTTAATAATCTTTGTCCAGAGACCAAGC 1740
Db 116403 GTGCGGCAAGTCAACATGATGGGGAGAAAGTTAATAATCTTTGTCCAGAGACCAAGC 116344
QY 1741 ACCCATCAACATCTGTGTAGTCTTAGAGGGCGAGAGAGGCTGTAGTGGGAGCTGG 1800
Db 116343 ACCCATCAACATCTGTGTAGTCTTAGAGGGCGAGAGAGGCTGTAGTGGGAGCTGG 116284
QY 1801 GAGGCTTTTCCCAAGAGGTGGCCTGTGAGCAGAGGCTCGGAAGATGACAGGTTTGACAGA 1860
Db 116283 GAGGCTTTTCCCAAGAGGTGGCCTGTGAGCAGAGGCTCGGAAGATGACAGGTTTGACAGA 116224
QY 1861 TGGGAAGTGGGGATGAGAGGACAGACGAGTGTTCAGGCCAAGGAACTGGGAACAAAGA 1920
Db 116223 TGGGAAGTGGGGATGAGAGGACAGACGAGTGTTCAGGCCAAGGAACTGGGAACAAAGA 116164
QY 1921 AGAACCTGAGAATGTAATCTACTTCAACCTGGACCTCTTGGCAGGGCTGCAATC 1980
Db 116163 AGAACCTGAGAATGTAATCTACTTCAACCTGGACCTCTTGGCAGGGCTGCAATC 116104
QY 1981 TCAGATGCCCTGAATGTGTGAAGTAGGCGGT 2011
Db 116103 TCAGATGCCCTGAATGTGTGAAGTAGGCGGT 116073

RESULT 3

AY436326

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

source

repeat_region

misc_feature

repeat_region

misc_feature

misc_feature

gene

mRNA

AY436326 23424 bp DNA linear PRI 26-OCT-2003
Homo sapiens angiotensin I converting enzyme (peptidyl-dipeptidase
A) 1 (ACE) gene, complete cds.

AY436326

AY436326.1 GI:37790803

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 23424)

Rieder, M.J., da Ponte, S.H., Kuldane, S.A., Rajkumar, N., Smith, J.D.,

Toth, E.J., Krauss, R.M. and Nickerson, D.A.

Direct Submission

Submitted (10-OCT-2003) Genome Sciences, University of Washington,

1705 NE Pacific, Seattle, WA 98195, USA

Location/Qualifiers

1..23424

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

54..205

/rpt_family="L2"

/rpt_type="dispersed"

1024..1143

/note="Region not scanned for variation"

1502..1593

/rpt_family="MIR"

/rpt_type="dispersed"

1688..1942

/note="Region not scanned for variation"

2097..2357

/note="Region not scanned for variation"

<2355..22724

/gene="ACE"

Join(<2355..2624,3212..3379,4288..4381,5050..5193,

5618..5809,6372..6469,6847..7019,7747..7970,8310..8454,

variation	/frequency="0.17"	8581	QY	781	CAGCATGTGGCCCGAGGCGGGGACTCTGTAAAGCCACTGCTGGAGAGCCATCTCCCATCCT	840
	/replaces="t"		Db	13724	CAGCATGTGGCCCGAGGCGGGGACTCTGTAAAGCCACTGCTGGAGAGCCATCTCCCATCCT	13783
	/gene="ACE"					
	/frequency="0.03"					
repeat_region	/replaces="t"		QY	841	TTCTCCCATTTTCTCTAGACCTGTGCCTATACAGTCACTTTTATGTGGTTTCGCCAATTT	900
	8608. .8650		Db	13784	TTCTCCCATTTTCTCTAGACCTGTGCCTATACAGTCACTTTTATGTGGTTTCGCCAATTT	13843
variation	/rpt_family="L2"					
	/rpt_type=dispersed					
	8683		QY	901	TATTCCAGCTCTGAAATTTCTCTGAGCTCCCTTAAAGCAGAGGTGAGCTAAGGGCTGGA	960
	/gene="ACE"		Db	13844	TATTCCAGCTCTGAAATTTCTCTGAGCTCCCTTAAAGCAGAGGTGAGCTAAGGGCTGGA	13903
	/frequency="0.43"					
variation	/replaces="t"		QY	961	GCTCAAGGCATTCAAAACCCCTACAGATCTGACGAATGTGATGGCCACGTCCTCCGGAATA	1020
	8824		Db	13904	GCTCAAGGCATTCAAAACCCCTACAGATCTGACGAATGTGATGGCCACGTCCTCCGGAATA	13963
	/gene="ACE"					
	/frequency="0.01"		QY	1021	TGAAGACCTGTTATGGGCAATGGAGGGCTGGGAGACAAGGCGGGGAGAGCCATCTCCCA	1080
	/replace="c"		Db	13964	TGAAGACCTGTTATGGGCAATGGAGGGCTGGGAGACAAGGCGGGGAGAGCCATCTCCCA	14023
Query Match	99.3%; Score 1997.4; DB 9; Length 23424;					
Best Local Similarity	99.9%; Pred. No. 0;					
Matches 2009; Conservative	0; Mismatches 1; Indels 1; Gaps 1;					
QY	1	CCGCGCCGGTTTGAAGAGCTCACCCCGACACAAGACCCGACACAGATACCTCCGAGC	60			
Db	12944	CCAGCCCGTTTGAAGAGCTCACCCCGACACAAGGACCCGACACAGATACCTCCGAGC	13003			
QY	61	TCCTCTCAACCCACCCCTTTTCAGGGTTGGAGAACTTGAGGCATAAATCTTGCTTCATGA	120			
Db	13004	TCCTCTCAACCCACCCCTTTTCAGGGTTGGAGAACTTGAGGCATAAATCTTGCTTCATGA	13063			
QY	121	GGAATCTCCACCCAGAAATGGGTCTTTCTGGCCCCCAGCCAGCTCCCAATTTAGAACAA	180			
Db	13064	GGAATCTCCACCCAGAAATGGGTCTTTCTGGCCCCCAGCCAGCTCCCAATTTAGAACAA	13123			
QY	181	TGACAAATAGAGGGGAAATGGAATAAACAAGGAGAAACGGTTTTCCAGGACAGGGTT	240			
Db	13124	TGACAAATAGAGGGGAAATGGAATAAACAAGGAGAAACGGTTTTCCAGGACAGGGTT	13183			
QY	241	TGGCCTACAAAGTTTGGATGTGGGTACCAATGCCAAGTGTGAGGGAGGCTGGCCGGGTG	300			
Db	13184	TGGCCTACAAAGTTTGGATGTGGGTACCAATGCCAAGTGTGAGGGAGGCTGGCCGGGTG	13243			
QY	301	TGTTGGCTCATGCTCTAATCCAGCACCTTTGGAGGCCCCAAGGTGATGATCACTTTGAG	360			
Db	13244	TGTTGGCTCATGCTCTAATCCAGCACCTTTGGAGGCCCCAAGGTGATGATCACTTTGAG	13303			
QY	361	GCGGGAGTTTGAGACAGCTGGGCCAACAATGSGTGAACCCCATCTGTACTAAAAATACA	420			
Db	13304	GCGGGAGTTTGAGACAGCTGGGCCAACAATGSGTGAACCCCATCTGTACTAAAAATACA	13363			
QY	421	AAAGTTAGCTGGGCGTGGTGTAGATGCCCTGTAGTCCAGCTACTTGGGAGGCTGAGGCA	480			
Db	13364	AAAGTTAGCTGGGCGTGGTGTAGATGCCCTGTAGTCCAGCTACTTGGGAGGCTGAGGCA	13423			
QY	481	TGAGAATCGCTTGAGCCAGCTGGGCAATACAGCAAGACCCCGTCTCTACAAATAAAAT	540			
Db	13424	TGAGAATCGCTTGAGCCAGCTGGGCAATACAGCAAGACCCCGTCTCTACAAATAAAAT	13483			
QY	541	ACAAAAATTTAGTTGATGTGGTGTGCATGTGCTGTAGTCTTGTAGGAGGCTGA	600			
Db	13484	ACAAAAATTTAGTTGATGTGGTGTGCATGTGCTGTAGTCTTGTAGGAGGCTGA	13543			
QY	601	GATGGAAGATTCTTTGAGCTGGGAGGTCAGGCTGCAGTGCAGATGGCGCCACT	660			
Db	13544	GATGGAAGATTCTTTGAGCTGGGAGGTCAGGCTGCAGTGCAGATGGCGCCACT	13603			
QY	661	GCACTCCAGCTGGGCAACAGAGTGAAGCCCTGCTCAGAAAAAATAAAAAAATAA	720			
Db	13604	GCACTCCAGCTGGGCAACAGAGTGAAGCCCTGCTCAGAAAAAATAAAAAAATAA	13663			
QY	721	GGAGAGGAGAGACTCAAGCACGCCCTTCACAGGACTGCTGAGGGCCCTGCAGGTGTCTG	780			
Db	13664	GGAGAGGAGAGACTCAAGCACGCCCTTCACAGGACTGCTGAGGGCCCTGCAGGTGTCTG	13723			

```
QY 1861 TGGGAAGTGGGGATGAGAGGACACGACGAGTGTTCAGGCCAAGGGAAGTGAACAAGA 1920
Db 14803 TGGGAAGTGGGGATGAGAGGACACGACGAGTGTTCAGGCCAAGGGAAGTGAACAAGA 14862
QY 1921 AGAAGCTGAGATGAAATCTACTTCAACCCCTGGACCCCTCCTTCCAGGGCTCAATC 1980
Db 14863 AGAAGCTGAGATGAAATCTACTTCAACCCCTGGACCCCTCCTTCCAGGGCTCAATC 14922
QY 1981 TCAGATGCCCTGAATGTGTGAAGTAGGCGGT 2011
Db 14923 TCAGATGCCCTGAATGTGTGAAGTAGGCGGT 14953

RESULT 4
AF118569
LOCUS AF118569 24070 bp DNA linear PRI 08-AUG-2000
DEFINITION Homo sapiens angiotensin I converting enzyme precursor (DCP1) gene,
alternative splice products, complete cds.
ACCESSION AF118569
VERSION AF118569.1 GI:4732025
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 24070)
AUTHORS Rieder,M.J., Taylor,S.L., Clark,A.G. and Nickerson,D.A.
TITLE Sequence variation in the human angiotensin converting enzyme
JOURNAL Nat. Genet. 22 (1), 59-62 (1999)
MEDLINE 99251580
PUBMED 10319862
REFERENCE 2 (bases 1 to 24070)
AUTHORS Rieder,M.J.
TITLE Direct Submission
JOURNAL Submitted (07-JAN-1999) Molecular Biotechnology, University of
Washington, UW Health Sciences Building, Rm. K316, 1705 NE Pacific,
Seattle, WA 98195, USA

FEATURES
source
1..24070
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="17"
/map="17q23 between D17S794 and D17S795"
variation
117
/frequency="0.05"
/replacement="t"
variation
334
/frequency="0.09"
/replacement="t"
repeat_region
complement(<337..>488)
/feature="putative"
/rpt_family="L2"
/rpt_type="dispersed"
repeat_region
<1785..>1841
/feature="putative"
/rpt_family="MIR"
/rpt_type="dispersed"
variation
2400
/frequency="0.32"
/replacement="t"
variation
2547
/frequency="0.23"
/replacement="c"
gene
<2662..>23217
/feature="DCP1"
mRNA
join(<2662..2010,3498..3665,4574..4667,5336..5479,
5904..6095,6658..6755,7133..7305,8032..8255,8595..8739,
9026..9124,9415..9537,9896..10107,10802..10938,
12123..12281,12552..12639,14499..14642,14792..14983,
16805..16902,17060..17232,19287..19510,19773..19917,
20223..20321,22245..22367,22649..22836,22988..>23217)
/feature="DCP1"
```

CDS

```
/product="angiotensin I converting enzyme precursor"
join(2662..2910,3498..3665,4574..4667,5336..5479,
5904..6095,6658..6755,7133..7305,8032..8255,8595..8739,
9026..9124,9415..9537,9896..10107,10802..10938,
12123..12281,12552..12639,14499..14642,14792..14983,
16805..16902,17060..17232,19287..19510,19773..19917,
20223..20321,22245..22367,22649..22836,22988..23217)
/feature="DCP1"
/feature="somatic form of ACE; dipeptidyl carboxypeptidase 1;
kininase II"
/codon_start=1
/product="angiotensin I converting enzyme precursor"
/protein_id="AAD28560.1"
/db_xref="GI:4732026"
/translation="MGAASRRGPGCLLLPLPLLLPPQALADPGLQPGNFSADEA
GAOLFAOSYNSAEOLVFSVAASWAHDTNLTAEANRQEEAALLSQFPAWGWAK
ELYPIMONFTDQPLRRTIGAVRTIGSANLPLAKHQYNALLSNMSRYSTAKVCLPN
KTATCWSLDPLTNILASSRSYMLLPFWGWHNRAGIPLPLEYDFDALSNEAYKQD
GFTDTGAYRWSYNSPTFDDLEHLYQLEPLYLNLHAFVRRALHRRYGDRIYINLRGP
IPALLGDMWAQSWENIYDMVVPFDPKPNLDVTSMLOQGNATHMFRVABEFFTSLE
LSPMPPEFWGSMLEKPADGREVVCASANDFYNRKDFRIKQCTRTVMDQSTVHHEM
GHIOYLOKDPVSLRGANPGFHEAIGDVALSVSTPEHLKIGLLDRVTNDTESD
INYLKMALEKIAFLPGYLDVDMWEGVSGRTPSRYNFDWYLRKYVQGCPPVTR
NETHFDAGAXPHVENVTPIRYEYFVLOFQFHEALCKEAGVEGHLQCDIYRSTKAG
AKLRVLAQSSRPQEWELKMDVGLDADLADQPLLYFQFVQMLQEQNQGEVLGWP
EYQWHPDLPNYPEGIDLVDEAEASKEVEEDRTSQVWNEYAEANWYNNTNITET
SKILLQKMQIANHTLYKGTQARFQVNLQNTTKRIKKVQDLERAAALPAQELEY
NKLIDMETTYSVATVCHPNQSCLEPDLTNWATSRKEDLLMAGWGLRDKAGRAI
LOFYPKVELINQARLYVDAGDSWRSMTYPSLEODLRLFOPLQPLNLHAYV
PRALHRYGAOHINLEGPIDPAHLGLNMAQTMNLYDLAVPPPSAAMDTEAMLKQG
WTPRRMFKAEADFTSLGLLPVPEFWNKSMLKPTDREVVVCHASANDFYNGKDFRI
KQCTTNLELDVAAHHEMGHIOYFMQYKDLVREGANPGFHEAIGDVALSVSTPK
HLNLSNLSEGGDEHDINFLKMDLKIAPFISYLVQDQWRVFDGSKTENYQ
EWSRLKYQGLCPVPRTQDFDPAKFIHPSVYRIFVFSFIQFHEALQAA
GHTGLKCDIYQSKAQORLATMKLGFSPWPEAMOLITGQPNMSASAMLSYFKPL
LDWRTENHLEHGLKGLWPQVNTNSASSEGLPDGSRVSLGLDLDAQARVQWLL
LFLGLALLVAILGLSQRFLSRHSLHSHSGPQPGSEVELRHS"
complement(2694..2737)
/feature="putative"
/rpt_type="tandem"
/rpt_unit="cag"
3872
/feature="DCP1"
/frequency="0.32"
/replacement="a"
<3951..>4250
/feature="putative"
/rpt_family="Alu"
/rpt_type="dispersed"
4230
/feature="DCP1"
/frequency="0.05"
/replacement="c"
4504
/feature="DCP1"
/frequency="0.41"
/replacement="c"
4733
/feature="DCP1"
/frequency="0.05"
/replacement="a"
4760
/feature="DCP1"
/frequency="0.05"
/replacement="c"
5406
/feature="DCP1"
/frequency="0.14"
/replacement="t"
5424
/feature="DCP1"
/frequency="0.05"
/replacement="a"
```

repeat_region

variation

repeat_region

variation

variation

variation

variation

variation

variation

variation	5492	/gene="DCP1"	QY	1	CCGGCCCGCTTTGAAAGAGCTCACCCCGACACAAAGACCGGCACACAGATACCTCCAGC	60
		/frequency="0.09"				
variation	5878	/replace="t"	Db	13228	CCAGCCCGTTTGAAGAGCTCACCCCGACACAAAGACCGGCACACAGATACCTCCAGC	13287
		/gene="DCP1"				
variation	6028	/frequency="0.09"	QY	61	TCCCTCTCAACCCACCCCTTTCCAGGGTTGGAGAACTTGAGGCATAAACTTGTCTCCATGA	120
		/replace="c"	Db	13288	TCCCTCTCAACCCACCCCTTTCCAGGGTTGGAGAACTTGAGGCATAAACTTGTCTCCATGA	13347
		/gene="DCP1"				
variation	6029	/frequency="0.05"	QY	121	GGAACTCTCACCCAGAAATGGGTCTTTCTGGCCCCAGCCCCAGCTCCCACTTAGAACAA	180
		/replace="t"	Db	13348	GGAACTCTCACCCAGAAATGGGTCTTTCTGGCCCCAGCCCCAGCTCCCACTTAGAACAA	13407
		/gene="DCP1"				
variation	6145	/frequency="0.05"	QY	181	TGACAAATAGAGGGGAATGAAATAAACAGGAGAAACGGTTTCCAGGACAGGGTT	240
		/replace="t"	Db	13408	TGACAAATAGAGGGGAATGAAATAAACAGGAGAAACGGTTTCCAGGACAGGGTT	13467
		/gene="DCP1"				
repeat_region	6320	/frequency="0.05"	QY	241	TGGCCTTACAAAGTTGTGGATGTGGGTACCCATGCCAAGTGTGAGGGAGGCTGGCCGGGTG	300
		/replace="t"	Db	13468	TGGCCTTACAAAGTTGTGGATGTGGGTACCCATGCCAAGTGTGAGGGAGGCTGGCCGGGTG	13527
		complement (<6245. .>6320)				
		/note="putative"	QY	301	TGGTGGCTCATGCTCTTAATCCAGACATTTGGGAGGCCAAGGTGAGTAGATCACTTGAG	360
		/rpt_family="MIR"	Db	13528	TGGTGGCTCATGCTCTTAATCCAGACATTTGGGAGGCCAAGGTGAGTAGATCACTTGAG	13587
		/rpt_type=dispersed				
variation	6435	/gene="DCP1"	QY	361	GCCGGGAGTTTGAGACCCAGCCTGGCCCAACATGGTGAACCCCATCTGTACTAAAAATACA	420
		/frequency="0.32"	Db	13588	GCCGGGAGTTTGAGACCCAGCCTGGCCCAACATGGTGAACCCCATCTGTACTAAAAATACA	13647
		/replace="g"				
variation	6604	/gene="DCP1"	QY	421	AAAGTTAGCTGGGCGTGGTGGTAGATGCCTGTAGTCCCAGCTACTTTGGGAGGCTGAGGCA	480
		/frequency="0.05"	Db	13648	AAAGTTAGCTGGGCGTGGTGGTAGATGCCTGTAGTCCCAGCTACTTTGGGAGGCTGAGGCA	13707
		/replace="a"				
repeat_region	7889	/frequency="0.05"	QY	481	TGAGATCGCTTTGAGCCAGCCTGGGCAATAACAGCAAGACCCCGCTCTTACAAATAAAT	540
		/note="putative"	Db	13708	TGAGATCGCTTTGAGCCAGCCTGGGCAATAACAGCAAGACCCCGCTCTTACAAATAAAT	13767
		/rpt_family="Alu"				
		/rpt_type=dispersed	QY	541	ACAAAAATTTAGTTGGATGTGGTGCATGCTGTAGTCTCTAGTCTGCTGAGGAGCTGA	600
		/gene="DCP1"	Db	13768	ACAAAAATTTAGTTGGATGTGGTGCATGCTGTAGTCTCTAGTCTGCTGAGGAGCTGA	13827
		/frequency="0.05"				
variation	7831	/replace="a"	QY	601	GATGAAGGATTGCTTGAGCCTGGGAGGTCAAGGCTGCAGTGCAGCGAGATGCGGCACCT	660
		/frequency="0.23"	Db	13828	GATGAAGGATTGCTTGAGCCTGGGAGGTCAAGGCTGCAGTGCAGCGAGATGCGGCACCT	13887
		/replace="a"				
variation	8128	/gene="DCP1"	QY	661	GCACTCCAGCCTGGGCAACAGAGTGAGACCCCTGTCTCAGAAAAAATAAAAAA	720
		/frequency="0.18"	Db	13888	GCACTCCAGCCTGGGCAACAGAGTGAGACCCCTGTCTCAGAAAAAATAAAAAA	13945
		/replace="t"				
repeat_region	8893	/note="putative"	QY	721	GGAGAGGAGAGACTCAAGCACGCCCTCAGAGACTGCTGAGGCCCTGCGAGGTCTCTG	780
		/rpt_family="L2"	Db	13946	GGAGAGGAGAGACTCAAGCACGCCCTCAGAGACTGCTGAGGCCCTGCGAGGTCTCTG	14005
		/rpt_type=dispersed				
variation	8935	/gene="DCP1"	QY	781	CAGCATGTGGCCCCAGGGCGGGGACTCTGTAAAGCCACTGCTGGAGCCACTCCCATCT	840
		/frequency="0.05"	Db	14006	CAGCATGT-GCCCCAGGGCGGGGACTCTGTAAAGCCACTGCTGGAG-CCACTCCCATCT	14063
		/replace="c"				
variation	8968	/gene="DCP1"	QY	841	TTCTCCCATTTCTTAGACCTGCTGCCT-----	868
		/frequency="0.36"	Db	14064	TTCTCCCATTTCTTAGACCTGCTGCCTATACAGTCACTTTTTTTTTTTTGTGAGACGG	14123
		/replace="t"				
variation	9191	/gene="DCP1"	QY	869	-----	868
		/frequency="0.09"	Db	14124	AGTCTCGTCTGTGCCCCAGGCTGGAGTGCAGTGGCGGGAATCTCGGCTCACTGCAAGCTC	14183
		/replace="t"				
variation	9222	/gene="DCP1"	QY	869	-----	868
		/frequency="0.09"	Db	14184	CGCTCCCGGGTTTACGGCCATTCTCTGCCTCAGCTCCCAAGTAGCTGGGACACAGGC	14243
		/replace="t"				
variation	9509	/gene="DCP1"	QY	869	-----	868
		/frequency="0.09"	Db	14244	GCCCCCCTACGCCCCGGCTAAATTTTTTTTGTATTTTTTAGTAGACGGGTTTTCACCGTTT	14303
		/replace="t"				

Query Match 82.9%; Score 1667.6; DB 9; Length 24070;
Best Local Similarity 87.1%; Pred. No. 0;
Matches 2003; Conservative 4; Indels 293; Gaps 4;

869 ----- 868
14304 TAGCGGATGGTCTCGATCTCTGACCTCGTATCCGCCGCTCGGCCTCCCAAAGTG 14363
869 -----ATACAGTCACCTTTATGTGTTTCGCCAATTTTATTCAGCTC 911
14364 CTGGGATTTACAGGCGTGATACAGTCACCTTTATGTGGTTTCGCCAATTTTATTCAGCTC 14423
912 TGAATTTCTCTGAGTCCCTTTACAGCAGAGGTGAGCTAAGGGCTGGAGCTCAAGGCAT 971
14424 TGAATTTCTCTGAGTCCCTTTACAGCAGAGGTGAGCTAAGGGCTGGAGCTCAAGGCAT 14483
972 TCAAACTCCCTACAGATCTGAGAAATGTATGGCCACGCTCCCGGAAATATGAAGCCTGT 1031
14484 TCAAACTCCCTACAGATCTGAGAAATGTATGGCCACGCTCCCGGAAATATGAAGCCTGT 14543
1032 TATGGCATGGAGGCTGGGAGACAGAGGGGGGAGAGCCATCTCCAGTTTATACCGA 1091
14544 TATGGCATGGAGGCTGGGAGACAGAGGGGGGAGAGCCATCTCCAGTTTATACCGA 14603
1092 AATAGTGGAACTCATCAACAGAGCTGCCGGCTCAATGGTGAGTCCCTGCTGCCAAT 1151
14604 AATAGTGGAACTCATCAACAGAGCTGCCGGCTCAATGGTGAGTCCCTGCTGCCAAT 14663
1152 CACTGGCACTGGGTCCTTCAATTTCTCAAGAGGTGCTGTAAGCCCAAGCCTAGG 1211
14664 CACTGGCACTGGGTCCTTCAATTTCTCAAGAGGTGCTGTAAGCCCAAGCCTAGG 14723
1212 AAAAGTATGATCCCTGGAGGAGGAGTAACTGTGTGTTGGAGAGCTGGCTGTGCC 1271
14724 AAAAGTATGATCCCTGGAGGAGGAGTAACTGTGTGTTGGAGAGCTGGCTGTGCC 14783
1272 CTCTGATGCTATGTAGATGACAGGAGCTCGTGAGGCTATGTACGAGACACCATCCCT 1331
14784 CTCTGATGCTATGTAGATGACAGGAGCTCGTGAGGCTATGTACGAGACACCATCCCT 14843
1332 GGAGCAAGACTGGAGCGGCTCTCCAGAGCTGAGGAGCTGACCTACCTGATGCTGATG 1391
14844 GGAGCAAGACTGGAGCGGCTCTCCAGAGCTGAGGAGCTGACCTACCTGATGCTGATG 14903
1392 CTACGTGGCGGGGCTGACCGTCACTACGGGGCCAGCAGCATCACTGGAGGGGCT 1451
14904 CTACGTGGCGGGGCTGACCGTCACTACGGGGCCAGCAGCATCACTGGAGGGGCT 14963
1452 CATTCCTGCTCACTGCTGGTAAAGGACATGTCGGGCTTTGAGAGGTTAAAGACGGA 1511
14964 CATTCCTGCTCACTGCTGGTAAAGGACATGTCGGGCTTTGAGAGGTTAAAGACGGA 15023
1512 CCAGAGTGTAGTGGGTTGGGACAGGCTGACTAGAGGTAGGAGCAGGCTGGGAC 1571
15024 CCAGAGTGTAGTGGGTTGGGACAGGCTGACTAGAGGTAGGAGCAGGCTGGGAC 15083
1572 TCAGAGACTCCAGCCTGTTGGGGGATGTTGCCAGGCTGGAGGGGCTGGGCTGGGA 1631
15084 TGAGAGACTCCAGCCTGTTGGGGATGTTGCCAGGCTGGAGGGGCTGGGCTGGGA 15143
1632 GTGGGAGCCCCCACTGTCATCTGGTGCCACATTCACCTGACAGTCTATGTCGGCAAGT 1691
15144 GTGGGAGCCCCCACTGTCATCTGGTGCCACATTCACCTGACAGTCTATGTCGGCAAGT 15203
1692 CACCATGATGGGGAGAGTAAATATCTGTCAGAGAGACCGGACCCATCACAA 1751
15204 CACCATGATGGGGAGAGTAAATATCTGTCAGAGAGACCGGACCCATCACAA 15263
1752 CATGTGTGATCTTACAGGGGAGAGAGGCTGTGAGTGGGAGCTGGGAGGCTTTGCC 1811
15264 CATGTGTGATCTTACAGGGGAGAGAGGCTGTGAGTGGGAGCTGGGAGGCTTTGCC 15323
1812 AAGAGTGGCTGTGAGAGGCGCTCGGAAGATGACAGGGTTTGAAGATGGGAAGTGGG 1871
15324 AAGAGTGGCTGTGAGAGGCGCTCGGAAGATGACAGGGTTTGAAGATGGGAAGTGGG 15383
1872 GGATGAGAGACAGACGCGAGTGTTCAGGCCAGGGGACCTGGACAAAGAACCTTGAGA 1931

15384 GGATGAGAGACAGACGCGAGTGTTCAGCCAGGAACTGGAAACAAAGAACCTTGAGA 15443
1932 ATGTAATCTACTTCAACCCCTGGACCTCTCTTTGCAAGGGCTGCAATCTCAGATSCCT 1991
15444 ATGTAATCTACTTCAACCCCTGGACCTCTCTTTGCAAGGGCTGCAATCTCAGATSCCT 15503
1992 GAATGTGTGAAGTAGGCGGT 2011
15504 GAATGTGTGAAGTAGGCGGT 15523
RESULT 5
A28005 Human ACE gene intron 15-18 sequence. linear PAT 18-SEP-1995
DEFINITION A28005
ACCESSION A28005
VERSION A28005.1 GI:1247488
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS
TITLE METHODS AND MEANS FOR STUDYING GENETIC POLYMORPHISM IN THE
ANGIOTENSIN CONVERTING ENZYME
JOURNAL Patent: WO 9300360-A 1 07-JAN-1993;
FEATURES
Location/Qualifiers
1. 2649
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 61.3%; Score 1233.6; DB 6; Length 2649;
Best Local Similarity 84.0%; Pred. No. 0;
Matches 1594; Conservative 0; Mismatches 295; Gaps 6;
QY 1 CCGGCCCTTTGAAAGAGCTCACCCCGACACAAAGAGCCGACACAGATACCTCCAGC 60
DB 759 CCAGCCCTTTGAAAGAGCTCACCCCGACACAAAGAGCCGACACAGATACCTCCAGC 818
QY 61 TCCCTCTCAACCCACCTTTCCAGGGTTGGAGAACTTGAGGCATAACTTGTCTCATGA 120
DB 819 TCCCTCTCAACCCACCTTTCCAGGGTTGGAGAACTTGAGGCATAACTTGTCTCATGA 878
QY 121 GGAATCTCCACCCAGAAATGGTCTTTCTGGCCCCCAGCCAGCTCCACATTAGAACAA 180
DB 879 GGAATCTCCACCCAGAAATGGTCTTTCTGGCCCCCAGCCAGCTCCACATTAGAACAA 938
QY 181 TGACAAATAGAGGGGAAATGAAATAAACAGAGAAACCGTTTTCAGGACAGGGTT 240
DB 939 TGACAAATAGAGGGGAAATGAAATAAACAGAGAAACCGTTTTCAGGACAGGGTT 998
QY 241 TGGCCCTCAAGTTGTGGATGTGGGTACCCATGCCAAGTGTGAGGGAGGCTGGCGGGTG 300
DB 999 TGGCCCTCAAGTTGTGGATGTGGGTACCCATGCCAAGTGTGAGGGAGGCTGGCGGGTG 1058
QY 301 TGGTGGCTCATGCTTAATCCAGCCTTTTGGAGGGCCAGAGGTGAGTAGATCACTTGA 360
DB 1059 TGGTGGCTCATG-CTCTAATCCAGCCTTTTGGAGGGCCAGAGGTGAGTAGATCACTTGA 1117
QY 361 GCCGGAGTTTGAAGCCAGCTTGGCCAAATAGTGGTGAACCCCATCTGTACTAAAAATACA 420
DB 1118 GCCGGAGTTTGAAGCCAGCTTGGCCAAATAGTGGTGAACCCCATCTGTACTAAAAATACA 1177
QY 421 AAAGTTAGTGGCGCTGTGGTAGATGCTGTAGTCCAGCTACTTGGAGGCTGAGGCA 480
DB 1178 AAAGTTAGTGGCGCTGTGGTAGATGCTGTAGTCCAGCTACTTGGAGGCTGAGGCA 1237
QY 481 TGAGAAATCGCTTGAGCCAGCCTGGGCAATACAGCAAGACCCCGCTCTTACAAATAAAT 540

Db 1238 TGAGATCGCTTGAGCCAGCCAGGGCAATACAGCAAGACCCCGTCTCTACAAATAAAT 1297
QY 541 ACAAAAATAGTGTGATGTGCTGATGCTGTAGTCTAGTCTAGGAGGCTGA 600
Db 1298 ACAAATAATAGTGTGATGTGCTGATGCTGTAGTCTAGTCTAGGAGGCTGA 1357
QY 601 GATGAAGGATTGCTTTGAGCTGGGAGGTCAAGGCTGCAGTGCAGCCGAGATGGCGCACT 660
Db 1358 GATGAAGGATTGCTTTGAGCTGGGAGGTCAAGGCTGCAGTGCAGCCGAGATGGCGCACT 1417
QY 661 GCATCTAGCTGGGCAACAGAGTGAGACCTGTCTCAGAAAAAATAAATAA 720
Db 1418 GCATCTAGCTGGGCAACAGAGTGAGACCTGTCTCAG-AAAAGAAAAAATAAATAA 1475
QY 721 GGAGAGGAGAGACTCAAGACGCCCCCTCACAGGACTGCTGAGGCCCTGCAGGCTGCTG 780
Db 1476 GGAGAGGAGAGACTCAAGACGCCCCCTCACAGGACTGCTGAGGCCCTGCAGGCTGCTG 1535
QY 781 CAGCATGTGGCCCCCAGGCCGGGACTCTGTAGCCACTGCTGGAGAGCCACTCCCATCT 840
Db 1536 CAGCATGTG-CCCAGGCCGGGACTCTGTAGCCACTGCTGGAGA-CCACTCCCATCT 1592
QY 841 TTCTCCCATTTCTTAGACTGTGCTCT----- 868
Db 1593 TTCTCCCATTTCTTAGACTGTGCTCTATACAGTCACTTTTTTTTTTTTGTGAGACGG 1652
QY 869 ----- 868
Db 1653 AGTCTCGCTGTGCGCCAGGCTGGAGTGAGTGGCGGAGTCTGGGCTCACTGCAACGTC 1712
QY 869 ----- 868
Db 1713 CGCCTCCGGGTTACGCCATTCTCTGCTCAGCTCAGCTCCCAAGTAGCTGGGACCAGCG 1772
QY 869 ----- 868
Db 1773 CCGGCCACTACGCCCGGCTAAATTTTTTTTATTTAGTAGAGAGCGGGTTTACCGTTTT 1832
QY 869 ----- 868
Db 1833 AGCCGGATGGTCTCGATCTCTGACTCTGATCGCTGATCGCCCGGCTCGGCTCCCAAGTGC 1892
QY 869 -----ATACAGTCACTTTTATGTGTTTTCGCCAATTTTATTTCCAGCTCT 912
Db 1893 TGGGATTACAGCGGTGATACAGTCACTTTTATGTGTTTTCGCCAATTTTATTTCCAGCTCT 1952
QY 913 GAAATTTCTAGTCTCCCTTACAGCAGAGGTGAGCTAAGGGCTGGAGCTCAAGGCATT 972
Db 1953 GAAATTTCTAGTCTCCCTTACAGCAGAGGTGAGCTAAGGGCTGGAGCTCAAGGCATT 2012
QY 973 CAAACCCCTTACCAGATCTGACGAATGTGATGCGCAGTCCCGGAAATATGAAGACTGTT 1032
Db 2013 CAAACCCCTTACCAGATCTGACGAATGTGATGCGCAGTCCCGGAAATATGAAGACTGTT 2072
QY 1033 ATGGGATGGGAGGCTGGCAGACAAGCGGGGAGCCATCTCAGTTTTTACCCGAA 1092
Db 2073 ATGGGATGGGAGGCTGGCAGACAAGCGGGGAGCCATCTCAGTTTTTACCCGAA 2132
QY 1093 ATACGTGNACTCATCAACAGGCTGCCGGCTCAATGTGAGTCCCTGCTGCCAATC 1152
Db 2133 ATACGTGNACTCATCAACAGGCTGCCGGCTCAATGTGAGTCCCTGCTGCCAATC 2192
QY 1153 ACTGGCAGTTGGTCCCTTCAATTTCTCAAGAGGTGCTGTGAACCCCAAGCCTAGGA 1212
Db 2193 ACTGGCAGTTGGTCCCTTCAATTTCTCAAGAGGTGCTGTGAACCCCAAGCCTAGGA 2252
QY 1213 AAAGGTAGATCCCTGGAGGAGCAGGTAAATGTGGTGTTCGGAGAGCCTGGCTGTGTCC 1272
Db 2253 AAAGGTAGATCCCTGGAGGAGCAGGTAAATGTGGTGTTCGGAGAGCCTGGCTGTGTCC 2312
QY 1273 TCTTAGGCTATGTAGATCGAGGAGCTCGTGGAGGTCTATGTAGAGACACCATCCCTG 1332
Db 2313 TCTTAGGCTATGTAGATCGAGGAGCTCGTGGAGGTCTATGTAGAGACACCATCCCTG 2372

QY 1333 GAGCAAGACTGAGCGGCTCTTTCAGGAGCTGCAGCCTACTTACTCAACCTGCATGCC 1392
Db 2373 GAGCAAGACTGAGCGGCTCTTTCAGGAGCTGCAGCCTACTTACTCAACCTGCATGCC 2432
QY 1393 TAGCTGCGCGGCGCTGCACCGTCACTACGGGGCCAGCAGCATCAACCTGGAGGGGCC 1452
Db 2433 TAGCTGCGCGGCGCTGCACCGTCACTACGGGGCCAGCAGCATCAACCTGGAGGGGCC 2492
QY 1453 ATTCTGTCTACCTGCTGGGTAAAGGCACATGTGCGGGCTTTGAGGAGGTAAAGACGGAC 1512
Db 2493 ATTCTGTCTACCTGCTGGGTAAAGGCACATGTGCGGGCTTTGAGGAGGTAAAGACGGAC 2552
QY 1513 CACAGTGTAGTGAGGTTGGGACAGGCTGACTAGAGGCTAGGAGCAGGCTGGGGACT 1572
Db 2553 CACAGTGTAGTGAGGTTGGGACAGGCTGACTAGAGGCTAGGAGCAGGCTGGGGACT 2612
QY 1573 GAGAGACTCCAGCCCTGTGGGGATGTTGCCAGGCT 1610
Db 2613 GAGAGACTCCAGCCCTGT-GGGATGTTGCCAGGCT 2649

RESULT 6
AR000113
LOCUS 1856 bp DNA linear PAT 04-DEC-1998
DEFINITION Sequence 3 from patent US 5736323.
ACCESSION AR000113
VERSION AR000113.1 GI:3962644
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1856)
AUTHORS Soubrier,F., Hubert,C. and Corvol,P.
TITLE Agents and procedures for the study of the genetic polymorphism of the angiotensin I converting enzyme
JOURNAL Patent: US 5736323-A 3 07-APR-1998,
FEATURES Location/Qualifiers
 source
 1..1856
 /organism="unknown"
 /mol_type="unassigned DNA"

ORIGIN
Query Match 41.0%; Score 824.6; DB 6; Length 1856;
Best Local Similarity 98.3%; Pred. No. 4.4e-223;
Matches 876; Conservative 0; Mismatches 9; Indels 6; Gaps 4;

QY 1 CCGGCCCGGTTTGAAGAGCTCACCCCGCACACAAGGACCCGACACAGATACCTCCAGC 60
Db 589 CCAGCCCGTTTGAAGAGCTCACCCCGCACACAAGGACCCGACACAGATACCTCCAGC 648
QY 61 TCCTCTCAACCCAGCTTTCCAGGTTGGAGAACTTGAAGGATAAACTTGTTCATGA 120
Db 649 TCCTCTCAACCCAGCTTTCCAGGTTGGAGAACTTGAAGGATAAACTTGTTCATGA 708
QY 121 GGAATCTCCACCCAGAAATGGTCTTCTGSCCCCGAGCCAGCTCCCATAGNACAA 180
Db 709 GGAATCTCCACCCAGAAATGGTCTTCTGSCCCCGAGCCAGCTCCCATAGNACAA 768
QY 181 TGACAAATAGAGGGGAAATGAAATAAACAAGAGGAAACGGTTTTCCAGGACAGGGTT 240
Db 769 TGACAAATAGAGGGGAAATGAAATAAACAAGAGGAAACGGTTTTCCAGGACAGGGTT 828
QY 241 TGGCCTCAAGTTGTGGATGGGTACCCATGCCCCAGTGTGAGGGAGGCTGGCCGGGTG 300
Db 829 TGGCCTCAAGTTGTGGATGGGTACCCATGCCCCAGTGTGAGGGAGGCTGGCCGGGTG 888
QY 301 TGGTGGCTCATGCTCTTAATCCAGCACTTTGGGAGGCCAAGGTGAGTAGATCACTTGA 360
Db 889 TGGTGGCTCATG-CTCTAATCCAGCACTTTGGGAGGCCAAGGTGAGTAGATCACTTGA 947
QY 361 GCGGGAGTTTGAAGACCAAGCTGGGCAACATGGTGTGAACCCCATCTGTACTAAATAACA 420

Db 948 GCCGGAGTTTGGAGCCAGCCTGGCCCAACATGGTGAAACCCCACTGTGTACTAAATAATACA 1007
 Qy 421 AAAGTTAGCTGGCGTGTGTGTAGATGCTGTAGTCCAGCTACTTGGAGAGCTTGAGGCA 480
 Db 1008 AAAGTTAGCTGGCGTGTGTGTAGATGCTGTAGTCCAGCTACTTGGAGAGCTTGAGGCA 1067
 Qy 481 TGAGAAATCGCTTGAGCCAGCCTGGGCAATACAGCAAGACCCCGTCTCTCAAAATAAAAT 540
 Db 1068 TGAGAAATCGCTTGAGCCAGCCTGGGCAATACAGCAAGACCCCGTCTCTCAAAATAAAAT 1127
 Qy 541 ACAAAAATTTAGTTGGATGTGGTGTGATGCTGTAGTCCAGCTACTTGGAGAGCTTGAGGCA 600
 Db 1128 ACAAAAATTTAGTTGGATGTGGTGTGATGCTGTAGTCCAGCTACTTGGAGAGCTTGAGGCA 1187
 Qy 601 GATGAAGGATTGCTTTGAGCCTGGGAGCTCAAGGCTCAAGGCTGAGCGGAGATGGCGCCACT 660
 Db 1188 GATGAAGGATTGCTTTGAGCCTGGGAGCTCAAGGCTCAAGGCTGAGCGGAGATGGCGCCACT 1247
 Qy 661 GCATCCAGCCTGGGCAACAGAGTGAGACCCCTGTCTAGAAATAAAAAAAAAAAAAA 720
 Db 1248 GCATCCAGCCTGGGCAACAGAGTGAGACCCCTGTCTAGAAATAAAAAAAAAAAAAA 1305
 Qy 721 GGAGGAGAGAGACTCAAGCAGCCCTCACAGAGCTGCTGAGGAGCCTGAGGCTGCTG 780
 Db 1306 GGAGGAGAGAGACTCAAGCAGCCCTCACAGAGCTGCTGAGGAGCCTGAGGCTGCTG 1365
 Qy 781 CAGCATGTGGCCCGCCAGGCGGGGACTCTGTAGCCACTGCTGAGGAGCCTGCCATCCT 840
 Db 1366 CAGCATGTG--CCAGGCGGGGACTCTGTAGCCACTGCTGAGGAGCCTGCCATCCT 1422
 Qy 841 TTCTCCCATTTCTTAGACCTGCTGCTATACAGTCACTTTTATGTGTTT 891
 Db 1423 TTCTCCCATTTCTTAGACCTGCTGCTATACAGTCACTTTTATGTGTTT 1473

RESULT 7
 AR137381
 LOCUS
 DEFINITION
 AR137381
 ACCESSION
 AR137381.1 GI:14478890
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Unknown.
 Unclassified.
 1 (bases 1 to 1856)
 Norberg,L.Torbjorn., Andersson,M.Kristina. and
 Lindstrom,P.Harry.Rutger.
 Methods for assessing cardiovascular status and compositions for
 use thereof
 Patent: US 6197505-A 128 06-MAR-2001;
 Location/Qualifiers
 1..1856
 /organism="unknown"
 /mol_type="unassigned DNA"

ORIGIN
 Query Match 41.0%; Score 824.6; DB 6; Length 1856;
 Best Local Similarity 98.3%; Pred. No. 4.4e-223;
 Matches 876; Conservative 0; Mismatches 9; Indels 6; Gaps 4;
 Qy 1 CCGGCCCGTTTGAAGAGCTACCCCGGACACAAGGACCCGACACAGATACCTCCAGC 60
 Db 589 CCAGCCCGTTTGAAGAGCTACCCCGGACACAAGGACCCGACACAGATACCTCCAGC 648
 Qy 61 TCCTCTCAACCCACCCCTTCCAGGTTTGGAGACTTGGAGGATTAACCTTCCATCA 120
 Db 649 TCCTCTCAACCCACCCCTTCCAGGTTTGGAGACTTGGAGGATTAACCTTCCATCA 708
 Qy 121 GGAATCTCCACCCAGAAATGGGTCTTTCTGGGCCCCCAGCCAGCTCCACATTAGAACAA 180
 Db 709 GGAATCTCCACCCAGAAATGGGTCTTTCTGGGCCCCCAGCCAGCTCCACATTAGAACAA 768

Qy 181 TGACAAATAGAGGGGAAATGGAATAAATAGAGAGAAACGGTTTTTCCAGACAGGTTT 240
 Db 769 TGACAAATAGAGGGGAAATGGAATAAATAGAGAGAAACGGTTTTTCCAGACAGGTTT 828
 Qy 241 TGGCCTACAGTTCTGGATGTGGGTAGCCATGCGCAAGTGTGAGGAGGCTGGCCGGGTG 300
 Db 829 TGGCCTACAGTTCTGGATGTGGGTAGCCATGCGCAAGTGTGAGGAGGCTGGCCGGGTG 888
 Qy 301 TGGTGGCTCATGCTCTTAATCCAGCACTTTGGGAGGCCAAGGTGAGTAGATCACTTCAG 360
 Db 889 TGGTGGCTCATG-CTCTAATCCAGCACTTTGGGAGGCCAAGGTGAGTAGATCACTTCAG 947
 Qy 361 GCCGGAGATTGGAGACAGCCTGGCCCAACATGTTGAAACCCCATCTGTACTAAATAATACA 420
 Db 948 GCCGGAGTTTGGAGACCACTGGCCCAACATGTTGAAACCCCATCTGTACTAAATAATACA 1007
 Qy 421 AAAGTTAGCTGGGCGTGTGTAGATGCTGTAGTCCAGCTACTTGGAGGCTGAGGCA 480
 Db 1008 AAAGTTAGCTGGGCGTGTGTAGATGCTGTAGTCCAGCTACTTGGAGGCTGAGGCA 1067
 Qy 481 TGAGAAATCGCTTGAGCCAGCCTGGGCAATACAGCAAGACCCCGTCTCTCAAAATAAAAT 540
 Db 1068 TGAGAAATCGCTTGAGCCAGCCTGGGCAATACAGCAAGACCCCGTCTCTCAAAATAAAAT 1127
 Qy 541 ACAAAAATTTAGTTGGATGTGGTGTGATGCTGTAGTCCAGCTACTTGGAGGCTGAGGCA 600
 Db 1128 ACAAAAATTTAGTTGGATGTGGTGTGATGCTGTAGTCCAGCTACTTGGAGGCTGAGGCA 1187
 Qy 601 GATGAAGGATTGCTTTGAGCCTGGGAGCTCAAGGCTCAAGGCTGAGCGGAGATGGCGCCACT 660
 Db 1188 GATGAAGGATTGCTTTGAGCCTGGGAGCTCAAGGCTCAAGGCTGAGCGGAGATGGCGCCACT 1247
 Qy 661 GCATCCAGCCTGGGCAACAGAGTGAGACCCCTGTCTAGAAATAAAAAAAAAAAAAA 720
 Db 1248 GCATCCAGCCTGGGCAACAGAGTGAGACCCCTGTCTAGAAATAAAAAAAAAAAAAA 1305
 Qy 721 GGAGGAGAGAGACTCAAGCAGCCCTCACAGAGCTGCTGAGGAGCCTGAGGCTGCTG 780
 Db 1306 GGAGGAGAGAGACTCAAGCAGCCCTCACAGAGCTGCTGAGGAGCCTGAGGCTGCTG 1365
 Qy 781 CAGCATGTGGCCCGCCAGGCGGGGACTCTGTAGCCACTGCTGAGGAGCCTGCCATCCT 840
 Db 1366 CAGCATGTG--CCAGGCGGGGACTCTGTAGCCACTGCTGAGGAGCCTGCCATCCT 1422
 Qy 841 TTCTCCCATTTCTTAGACCTGCTGCTATACAGTCACTTTTATGTGTTT 891
 Db 1423 TTCTCCCATTTCTTAGACCTGCTGCTATACAGTCACTTTTATGTGTTT 1473

RESULT 8
 BD231364
 LOCUS
 DEFINITION
 BD231364
 ACCESSION
 BD231364.1 GI:33041134
 VERSION
 KEYWORDS
 JP 2002527079-A/128.
 SOURCE
 Homo sapiens (human)
 ORGANISM
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini, Hominidae; Homo.
 1 (bases 1 to 1856)
 Norberg,L.T., Andersson,M.K., Lindstrom,P.H.R. and Jonsson,L.
 Genes for assessing cardiovascular status and compositions for use
 thereof
 Patent: JP 2002527079-A 128 27-AUG-2002;
 PAIROSEAKENSINGU AB
 OS Homo sapiens (human)
 PN JP 2002527079-A/128
 PD 27-AUG-2002
 PF 13-OCT-1999 JP 2000576056
 PR 14-OCT-1998 US 60/104286,14-OCT-1998 US 60/104302 PI
 LEIF TORBUORN NORBERG,MARIA KRISTINA ANDERSSON,PER HARRY PI


```
RUTGER LINDSTROM,
PI LENA JONSSON
PC C1201/68.C12N15/09//G01N33/53.G01N33/566.C12N15/00 CC Genes
for assessing cardiovascular status
and compositions for
CC use thereof
FH Location/Qualifiers
FT 1..1856
FT source /organism='Homo sapiens (human)'.

FEATURES
source
1..1856
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 41.0%; Score 824.6; DB 6; Length 1856;
Best Local Similarity 98.3%; Pred. No. 4.4e-223;
Matches 876; Conservative 0; Mismatches 9; Indels 6; Gaps 4;

QY 1 CCGGCCGCTTTGAAAGAGCTCACCCCGACACAAAGGACCGGCACAGATACCTCCAGC 60
Db 589 CAGCCCGCTTTGAAAGAGCTCACCCCGACACAAAGGACCGGCACAGATACCTCCAGC 648
QY 61 TCCCTCTCAACCCACCCCTTTCCAGGGTTGGAGAACTTGAGGCATAAACTTCTCCATGA 120
Db 649 TCCCTCTCAACCCACCCCTTTCCAGGGTTGGAGAACTTGAGGCATAAACTTCTCCATGA 708
QY 121 GGAAATCCACCCAGAAATGGGTCTTTCTGGCCCCCAGGCCAGCTCCCATTTAGAACAA 180
Db 709 GGAAATCCACCCAGAAATGGGTCTTTCTGGCCCCCAGGCCAGCTCCCATTTAGAACAA 768
QY 181 TGACAAATAGAGGGGAAATGGAATAAACAAGGAGAAACGGTTTCCAGGACAGGGTT 240
Db 769 TGACAAATAGAGGGGAAATGGAATAAACAAGGAGAAACGGTTTCCAGGACAGGGTT 828
QY 241 TGGCCCTACAAAGTTGGATGTGGGTACCCATCCAGTGTAGGGAGAGCTGCGCGGGTG 300
Db 829 TGGCCCTACAAAGTTGGATGTGGGTACCCATCCAGTGTAGGGAGAGCTGCGCGGGTG 888
QY 301 TGGTGGCTCATGCTCTAATCCAGCAGCTTTGGGAGGCGCAAGGTGAGTAGATCACTTGAG 360
Db 889 TGGTGGCTCATG-CTCTAATCCAGCAGCTTTGGGAGGCGCAAGGTGAGTAGATCACTTGAG 947
QY 361 GCGGGAGTTTGAGACGAGCTGGCCGACATGTTGAAACCCCATCTGTACTAAAAATACA 420
Db 948 GCGGGAGTTTGAGACGAGCTGGCCGACATGTTGAAACCCCATCTGTACTAAAAATACA 1007
QY 421 AAAGTTAGCTGGGCGTGGTGTAGATGCTGTAGTCCAGCTACTTGGGAGGCTGAGGCA 480
Db 1008 AAAGTTAGCTGGGCGTGGTGTAGATGCTGTAGTCCAGCTACTTGGGAGGCTGAGGCA 1067
QY 481 TGAGATCGCTTGAGCCAGCTGGGCAATACAGCAAGACCCCGTCTTACAAATAAAT 540
Db 1068 TGAGATCGCTTGAGCCAGCTGGGCAATACAGCAAGACCCCGTCTTACAAATAAAT 1127
QY 541 ACAAAAAATTTAGTTGATGTGGTGGCATGCTGTAGTCTAGTCTGCTAGGAGGCTGA 600
Db 1128 ACAAAAAATTTAGTTGATGTGGTGGCATGCTGTAGTCTAGTCTGCTAGGAGGCTGA 1187
QY 601 GATGAAGGATTTGCTTGGAGCTGGGAGGTCAAGGCTGCAGTGCAGCGAGATGGCGCACT 660
Db 1188 GATGAAGGATTTGCTTGGAGCTGGGAGGTCAAGGCTGCAGTGCAGCGAGATGGCGCACT 1247
QY 661 GCATCCAGCTGGGCAACAGAGTGAGACCTGTCTCAGAAAAAATAAAAAAATAAAAA 720
Db 1248 GCATCCAGCTGGGCAACAGAGTGAGACCTGTCTCAG-AAAAGAAAAAATAAAAAA 1305
QY 721 GGAGAGGAGAGACTCAAGACGCGCCCTCAGAGGACTGCTGAGGCGCTGCGAGGTCTG 780
Db 1306 GGAGAGGAGAGACTCAAGACGCGCCCTCAGAGGACTGCTGAGGCGCTGCGAGGTCTG 1365
QY 781 CAGCATGTGGCCCCAGGCGGGGACTCTGTGAAGCCACTGTGTGAGAGCCACTCCCATCT 840
```

```
Db 1366 CAGCATGTG-CCAGCGCGGGGACTCTGTAGCCACTGCTGGAGA-CCACTCCATCT 1422
QY 841 TTCTCCCATTTCTTAGACTGCTGCTATACAGTCACTTTTAATGTTT 891
Db 1423 TTCTCCCATTTCTTAGACTGCTGCTATACAGTCACTTTT 1473

RESULT 9
BD075256 1856 bp DNA linear PAT 27-AUG-2002
LOCUS Methods for assessing cardiovascular status and compositions for
DEFINITION use thereof.
ACCESSION BD075256
VERSION BD075256.1 GI:22620859
KEYWORDS JP 2001519660-A/129.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1856)
AUTHORS Norberg,L.T., Andersson,M.K. and Lindstrom,P.H.R.
TITLE Methods for assessing cardiovascular status and compositions for
use thereof
JOURNAL Patent: JP 2001519660-A 129 23-OCT-2001;
COMMENT EURONA MEDICAL AB
OS Homo sapiens (human)
PN JP 2001519660-A/129
PD 23-OCT-2001
PF 01-APR-1998 JP 1998542530
PI 04-APR-1997 US 60/042930
PI LEIF TORBJORN NORBERG,MARIA KRISTINA ANDERSSON,PER HARRY PI
RUTGER LINDSTROM
PC C1201/68.C07K14/72.C07K14/575.C12N9/48
CC ACE intron 16
FH Key Location/Qualifiers
FT source 1..1856
FT /organism='Homo sapiens (human)'.

FEATURES
source
1..1856
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 41.0%; Score 824.6; DB 6; Length 1856;
Best Local Similarity 98.3%; Pred. No. 4.4e-223;
Matches 876; Conservative 0; Mismatches 9; Indels 6; Gaps 4;

QY 1 CCGGCCGCTTTGAAAGAGCTCACCCCGACACAAAGGACCGGCACAGATACCTCCAGC 60
Db 589 CAGCCCGCTTTGAAAGAGCTCACCCCGACACAAAGGACCGGCACAGATACCTCCAGC 648
QY 61 TCCCTCTCAACCCACCCCTTTCCAGGGTTGGAGAACTTGAGGCATAAACTTCTCCATGA 120
Db 649 TCCCTCTCAACCCACCCCTTTCCAGGGTTGGAGAACTTGAGGCATAAACTTCTCCATGA 708
QY 121 GGAAATCCACCCAGAAATGGGTCTTTCTGGCCCCCAGGCCAGCTCCCATTTAGAACAA 180
Db 709 GGAAATCCACCCAGAAATGGGTCTTTCTGGCCCCCAGGCCAGCTCCCATTTAGAACAA 768
QY 181 TGACAAATAGAGGGGAAATGGAATAAACAAGGAGAAACGGTTTCCAGGACAGGGTT 240
Db 769 TGACAAATAGAGGGGAAATGGAATAAACAAGGAGAAACGGTTTCCAGGACAGGGTT 828
QY 241 TGGCCCTACAAAGTTGGATGTGGGTACCCATCCAGTGTAGGGAGAGCTGCGCGGGTG 300
Db 829 TGGCCCTACAAAGTTGGATGTGGGTACCCATCCAGTGTAGGGAGAGCTGCGCGGGTG 888
QY 301 TGGTGGCTCATGCTCTAATCCAGCAGCTTTGGGAGGCGCAAGGTGAGTAGATCACTTGAG 360
Db 889 TGGTGGCTCATG-CTCTAATCCAGCAGCTTTGGGAGGCGCAAGGTGAGTAGATCACTTGAG 947
```

Qy	361	GC	CGGAGTTTGAGACAGCGCTGCCCAACATGCTGAACCCCATCTGTACTAAAAATACA	420
Db	948	GC	CGGAGTTTGAGACAGCGCTGCCCAACATGCTGAACCCCATCTGTACTAAAAATACA	1007
Qy	421	AA	AGTTAGCTGGGGCGTGGTATAGTCCCTGTAGTCCACGCTACTCTGGGAGGCTGAGGCA	480
Db	1008	AA	AGTTAGCTGGGGCGTGGTATAGTCCCTGTAGTCCACGCTACTCTGGGAGGCTGAGGCA	1067
Qy	481	TG	GAATCGCTTGAGCCCGAGCTGGGCAATACAGCAAGACCCCGTCTCTCAAAATAAAAT	540
Db	1068	TG	GAATCGCTTGAGCCCGAGCCAGCGGCAATACAGCAAGACCCCGTCTCTCAAAATAAAAT	1127
Qy	541	AC	AAAAAATAGTTGGATGTGGTGCATGCTCCCTGTAGTCTTAGCTAGCGAGGCTGA	600
Db	1128	AC	AAAAAATAGTTGGATGTGGTGCATGCTCCCTGTAGTCTTAGCTAGCGAGGCTGA	1187
Qy	601	GA	TGGAAGGATGTGCTTGAGCCTCGGAGGGTCAAGGCTCGATGAGCCGAGATGGCGCACT	660
Db	1188	GA	TGGAAGGATGTGCTTGAGCCTCGGAGGGTCAAGGCTCGATGAGCCGAGATGGCGCACT	1247
Qy	661	GC	ACTCCAGCCTGGGCAACAGAGTGACACCTGTCTCAGAAAAAATAAAAAAAAAAAAA	720
Db	1248	GC	ACTCCAGCCTGGGCAACAGAGTGAGACCTGTCTCAG--AAAGAAAAAATAAAAAAAAA	1305
Qy	721	GG	GAGGAGAGAGACTCAAGCACGCCCTCTCACAGGATGTCTGAGGCCCTCGAGGTGCTG	780
Db	1306	GG	GAGGAGAGAGAGACTCAAGCACGCCCTCTCACAGGACTGTCTGAGGCCCTCGAGGTGCTG	1365
Qy	781	CAG	CATGTGGCCCAACAGCCGGGAGCTCTGTAAAGCCACTGCTGGAGAGGCCACTCCCATCCT	840
Db	1366	CAG	CATGTG--CCCAAGCCGGGAGCTCTGTAAAGCCACTGCTGGAGAGGCCACTCCCATCCT	1422
Qy	841	TT	TCCCATTTCTCTAGACCTGCTGCCCTATACAGTCACTTTTATGTGGTTT	891
Db	1423	TT	TCCCATTTCTCTAGACCTGCTGCCCTATACAGTCACTTTTATGTGGTTT	1473

```

RESULT 10
BD075259
LOCUS
DEFINITION
BD075259
Methods for assessing cardiovascular status and compositions for
use thereof.
ACCESSION
BD075259
GI:22620862
VERSION
JP 2001519660-A/132.
KEYWORDS
Homo sapiens (human)
SOURCE
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (Bases 1 to 1856)
AUTHORS
Norberg,L.T., Andersson,M.K. and Lindstrom,P.H.R.
TITLE
Methods for assessing cardiovascular status and compositions for
use thereof
JOURNAL
Patent: JP 2001519660-A 132 23-OCT-2001;
EUPONA MEDICAL AB
COMMENT
OS Homo sapiens (human)
PN JP 2001519660-A/132
PD 23-OCT-2001
PF 01-APR-1998 JP 1998542530
PI 04-APR-1997 US 60/042930
PR LEIF TORBJORN NORBERG, MARIA KRISTINA ANDERSSON, PER HARRY PI
RUTGER LINDSTROM
CC C12Q1/68, C07K14/72, C07K14/575, C12N9/48
CC Deletion
FH Key
FT source
FT Location/Qualifiers
1..1856
/organism='Homo sapiens (human)'.
FEATURES
source
1..1856
/organism='Homo sapiens'
/mol_type='genomic DNA'
/db_xref='taxon:9606'
ORIGIN

```

polymorphic variation.
 Homo sapiens (human)
 ORGANISM
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE
 1 (bases 1 to 1856)
 Rigat,B., Hubert,C., Corvol,P. and Soubrier,F.
 PCR detection of the insertion/deletion polymorphism of the human
 angiotensin converting enzyme gene (ACE1) (dipeptidyl
 carboxypeptidase 1)
 Nucleic Acids Res. 20 (6), 1433 (1992)
 JOURNAL
 MEDLINE
 92220641
 PUBMED
 1313972
 REFERENCE
 2 (bases 1 to 1856)
 Soubrier,F.
 Direct Submission
 TITLE
 Submitted (25-OCT-1991) F. Soubrier, Inserm Unit 36 College de
 France, 3 Rue d'Ulm, 75005 Paris, FRANCE
 JOURNAL
 COMMENT
 Related sequences: Hubert,C. et al, J.Biol.Chem. 266:15377-15383
 (1991).
 FEATURES
 Location/Qualifiers
 1..1856
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="17"
 /map="q23"
 /clone="Lambda 311"
 /tissue_type="placenta"
 /clone_lib="genomic"
 880..1105
 /note="incomplete Alu type sequence"
 repeat_region
 1106..1301
 /note="incomplete Alu type sequence"
 repeat_region
 1451..1738
 /note="insertion/deletion polymorphism due to
 presence/absence of sequence"
 complement(1451..1738)
 /note="Alu type sequence in antisense orientation"
 ORIGIN
 Query Match 41.0%; Score 824.6; DB 9; Length 1856;
 Best Local Similarity 98.3%; Pred. No. 4.4e-223;
 Matches 876; Conservative 0; Mismatches 9; Indels 6; Gaps 4;
 QY 1 CGGGCCCGTTGAAGAGCTCACCCCGCACAGGACCGCACAGATACCTCCAC 60
 Db 589 CCAGCCCGTTGAAGAGCTCACCCCGCACAGGACCGCACAGATACCTCCAC 648
 QY 61 TCCCTCTCAACCCACCCCTTTCCAGGGTTGGAGAACTTGAGGCATAAATCTTCCATGA 120
 Db 649 TCCCTCTCAACCCACCCCTTTCCAGGGTTGGAGAACTTGAGGCATAAATCTTCCATGA 708
 QY 121 GGAATCTCCACCCAGAAATGGGTCTTTCTGGCCCGCCAGCCAGCTCCCATAGAACAA 180
 Db 709 GGAATCTCCACCCAGAAATGGGTCTTTCTGGCCCGCCAGCCAGCTCCCATAGAACAA 768
 QY 181 TCACAAATAGAGGGGAAATGGAATAACAGGAGAAACGGTTTCCAGGACAGGGTT 240
 Db 769 TCACAAATAGAGGGGAAATGGAATAACAGGAGAAACGGTTTCCAGGACAGGGTT 828
 QY 241 TGGCCCTACAAGTTGTGGATGTGGGTACCCCATGCCAAGTGTGAGGGAGGCTGGCCGGGTG 300
 Db 829 TGGCCCTACAAGTTGTGGATGTGGGTACCCCATGCCAAGTGTGAGGGAGGCTGGCCGGGTG 888
 QY 301 TGGTGGCTCATGCTTAATCCAGCACTTTGGAGGGCCAAAGGTAGTAGTACCTTGAG 360
 Db 889 TGGTGGCTCATG-CTTAATCCAGCACTTTGGAGGGCCAAAGGTAGTAGTACCTTGAG 947
 QY 361 GCGGGAGTTTGAGACCGCTGGCCCAACATGTTGTAACCCCATCTGTACTAAATAACA 420
 Db 948 GCGGGAGTTTGAGACCGCTGGCCCAACATGTTGTAACCCCATCTGTACTAAATAACA 1007

QY 421 AAAGTTAGCTGGCGCTGGTGGTAGATGCTGTAGTCCAGCTACTTGGGAGGCTGAGGCA 480
 Db 1008 AAAGTTAGCTGGCGCTGGTGGTAGATGCTGTAGTCCAGCTACTTGGGAGGCTGAGGCA 1067
 QY 481 TGAGAATCGCTTTAGCCAGCCAGCTGGGCAATACAGCAAGACCCCGTCTCTACAAATAAAAT 540
 Db 1068 TGAGAATCGCTTTAGCCAGCCAGCTGGGCAATACAGCAAGACCCCGTCTCTACAAATAAAAT 1127
 QY 541 ACACAAAATTTAGTGGATGTGGTGCATGCTGTAGTCCCTAGTCTGCTAGGAGGCTGA 600
 Db 1128 ACACAAAATTTAGTGGATGTGGTGCATGCTGTAGTCCCTAGTCTGCTAGGAGGCTGA 1187
 QY 601 GATGAAGGATTGCTTGAGCTGGGAGGTCAAGGCTGAGTGCAGCCGATGCGGCCT 660
 Db 1188 GATGAAGGATTGCTTGAGCTGGGAGGTCAAGGCTGAGTGCAGTGCAGGAGTGGCGCCT 1247
 QY 661 GCATCCAGCTGGGCAACAGAGTGCAGCCCTGTCTCAGAAAAAATAAAAAA 720
 Db 1248 GCATCCAGCTGGGCAACAGAGTGCAGCCCTGTCTCAG--AAAGAAAAAATAAAAAA 1305
 QY 721 GGAGAGGAGAGACTCAAGCAGCCCTCAGAGACTGCTGAGCCCTGCGAGTGTCTG 780
 Db 1306 GGAGAGGAGAGACTCAAGCAGCCCTCAGAGACTGCTGAGCCCTGCGAGTGTCTG 1365
 QY 781 CAGCATGTGGCCCGGCGGAGCTCTGTAAGCCACTGCTGAGAGCCACTCCCATCT 840
 Db 1366 CAGCATGTG--CCAGGCGGCGGAGCTCTGTAAGCCACTGCTGAGAG--CACTCCCATCT 1422
 QY 841 TTCTCCCATTTCTTAGACCTGCTGCTATACAGTCACTTTTATGTGTTT 891
 Db 1423 TTCTCCCATTTCTTAGACCTGCTGCTATACAGTCACTTTTATGTGTTT 1473
 RESULT 12
 AC118572
 LOCUS
 DEFINITION
 Lemur catta clone LB2-253K19, WORKING DRAFT SEQUENCE, 12 ordered
 pieces.
 AC118572
 AC118572.2 GI:32528927
 HTG; HTGS PHASE2; HTGS DRAFT.
 SOURCE
 Lemur catta (ring-tailed lemur)
 ORGANISM
 Lemur catta
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Strepsirhini; Lemnidae; Lemur.
 REFERENCE
 1 (bases 1 to 195863)
 Cheng,J.-F., Hamilton,M., Peng,Y., Mukherjee,S., Hosseini,R.,
 Peng,Z., Malinov,I. and Rubin,E.M.
 Direct Submission
 TITLE
 Unpublished
 REFERENCE
 2 (bases 1 to 195863)
 Martin,J., Hosseini,R., Peng,Y., Peng,Z., Rubin,E.M. and
 Cheng,J.-F.
 Direct Submission
 JOURNAL
 Submitted (19-APR-2002) Genome Sciences, Lawrence Berkeley National
 Laboratory, 1 Cyclotron Rd., Berkeley, CA 94720, USA
 REFERENCE
 3 (bases 1 to 195863)
 Cheng,J.-F., Hamilton,M., Peng,Y., Mukherjee,S., Hosseini,R.,
 Peng,Z., Malinov,I. and Rubin,E.M.
 Direct Submission
 TITLE
 Submitted (11-JUL-2003) Genome Sciences, Lawrence Berkeley National
 Laboratory, 1 Cyclotron Rd., Berkeley, CA 94720, USA
 JOURNAL
 COMMENT
 On Jul 11, 2003 this sequence version replaced gi:20198536.

Sequence Produced by Berkeley PGA
 Web site: <http://pga.lbl.gov>
 Center Code: PGABERK
 Center Project Name: L027
 Bac Clone Name: LB2-253K19

This sequence has been compared to sequences of other species
 using Vista (<http://www-gsd.lbl.gov/VISTA>). The results can be
 viewed at:

http://pga.lbl.gov/cgi-bin/search_cvcgd?type=n&value=ACE

The order-orientation of the draft sequence was accomplished by using:

Avid (<http://baboon.math.berkeley.edu/mavid/>),
Lagan (<http://lagan.stanford.edu/>) and paired end information.

Funding agent: Programs for Genomic Applications (NHLBI)

If the Bac Library Name is LB1 to LB4, please see website for the description: <http://www-gsd.lbl.gov/cheng/BAC.html>
These libraries are available through the BACPAC Resources Center: <http://www.chori.org/bacpac/libraryres.htm> as LBNL-1 to LBNL-4.

Summary Statistics:

Sequencing vector: Plasmid; pUC18

Chemistry: Dye-terminator Big Dye

Assembly program: Phrap version 0.990329.

* NOTE: This is a 'working draft' sequence. It currently

* consists of 12 contigs. Gaps between the contigs

* are represented as runs of N. The order of the pieces

* is believed to be correct as given, however the sizes

* of the gaps between them are based on estimates that have

* provided by the submitter.

* This sequence will be replaced

* by the finished sequence as soon as it is available and

* the accession number will be preserved.

* 1 10817: contig of 10817 bp in length

* 10818 10917: gap of unknown length

* 10918 19222: contig of 8305 bp in length

* 19223 19322: gap of unknown length

* 19323 52285: contig of 32963 bp in length

* 52286 52385: gap of unknown length

* 52386 75793: contig of 23408 bp in length

* 75794 75893: gap of unknown length

* 75894 99277: contig of 23384 bp in length

* 99278 99377: gap of unknown length

* 99378 107543: contig of 8166 bp in length

* 107544 107643: gap of unknown length

* 107644 118968: contig of 11225 bp in length

* 118969 118968: gap of unknown length

* 118969 121200: contig of 2232 bp in length

* 121201 121300: gap of unknown length

* 121301 139834: contig of 18534 bp in length

* 139835 139934: gap of unknown length

* 139935 156316: contig of 16382 bp in length

* 156317 156416: gap of unknown length

* 156417 167012: contig of 10596 bp in length

* 167013 167112: gap of unknown length

* 167113 195863: contig of 28751 bp in length.

FEATURES

source

```
1. 195863
/organism="Lemur catta"
/mol_type="genomic DNA"
/db_xref="taxon:9447"
/clone="LB2-253K19"
```

ORIGIN

```
Query Match      24.4%; Score 490.6; DB 2; Length 195863;
Best Local Similarity 72.0%; Pred. No. 9.3e-128;
Matches 860; Conservative 0; Mismatches 264; Indels 71; Gaps 14;

QY      843 CTCCTATTTCTAGACCTGCTGCTATACAGTCACTTTATGTGGTTTCGCCAATTTTA 902
DB      128866 CCCCAGCTGCTGGCGCTTCTCCAGGGGCGAGAACTTTCAGTGGTCTCTGC-ATTGCA 128924

QY      903 TTCCAGCTCTGAATTTCTAGCTCCCTTACAGCGAG-----AGGTGAGCTAAGGGCT 957
DB      128925 CTCCTGCTCTGAGAGCTGTGAGTCCCTCGAAGAGAGGTTGGTGGCTGAGCGAAGGGCT 128984

QY      958 GGAGCTCAAGGATTCMAACCC-----CTACAGATCTCAAGATGTGATGCCAGG 1009
DB      128985 GGGGCTCAGCTGTTCGACCCCAACCCCAACCCCAACCTCCAGATCTGACAGCGTGTGCGCACA 129044
```

```
QY      1010 TCCGGAATATGAAGACCTGTTATGGGCATGGGAGGCTGGCGAGACAAGCGGGGAGA 1069
DB      129045 TCCGGAAGTACGAAGAACTGTTATGGGCTTGAAGGCTGGCGAGACCAAGTGGGAAGA 129104

QY      1070 GCCATCTCTCCAGTTTATACCCGAATACTGTAAGAACTCATCAACCAAGGCTGCCGCGCTCAAT 1129
DB      129105 GCCATCTCTCTCTTTTCCCAAAATACGTGAACTCACCAACAAAGGCTGCCCTGCTCAAC 129164

QY      1130 GGTGAGTCCCTGCTGCCAACAATCACTGCACTTGGGTCCCTTCATTTTCTCAAGAGGT 1189
DB      129165 GGTGAGTCCCAAGTAGCC-CTGGCTCTCTGGTCCCTTGGGCGCTTCCCTCAAGAGGT 129223

QY      1190 GCTGTGAACCCCAAGCCCTAGGAAAGGTAGATCCCTGGAGGAGCGAGTAA----- 1241
DB      129224 CCTCTGAGACCGTAAGTCTGGGGAAGG-AGGGCCCTGGAGGGGCGAGGTGAGGG 129282

QY      1242 ----TGTGGTGTGGGAGAGCTGGCTGTGTGCCCTCTGTAG-GCTATGTAGATCAGAG 1296
DB      129283 ATGGTCAGGTGTTCGAAGAGACTGGCTGTGTCTCCCTTCGAGAGCTACGCGAGCGAGGG 129342

QY      1297 GACTCGTGGAGTCTATGTACGAGACACCATCCCTGGAGCAAGACTCGAGCGGCTCTTC 1356
DB      129343 GATTCTGGAGGTCTATGTACGAGACCATCCCTGGAGCAAGACTCGAGCGAGTCTTAC 129402

QY      1357 CAGGAGCTGCAGCCACTCTACTCAACCTGCATGCTACGTGCGCGCGGCGCTGCACCGT 1416
DB      129403 CAGGAGCTGCAGCGCTTTACTGAACCTGCACGCTACGTGCGCGCGGCGCTGCACCGC 129462

QY      1417 CACTACGGGGCCCGAGCACATCAACTGAGGGGGCCATCTCTGTCTACCTGCTGGGTAG 1476
DB      129463 CACTACGGGGCCCGAGCACATCAACTGAGGGGGCCATCTCTGTCTCATCTGCTGGGTAG 129522

QY      1477 GGCACATCTCGGCTTTCAGGAGGTAAAGACGACACAGTGTGAGTGAGGTTGGGAC 1536
DB      129523 GGCACAGCTCGGCTTTCAGGAGGTAAAGACGACACAGTGTGAGTGAGGTTGGGAC 129581

QY      1537 AGGGCTGACTAGAGGTAGGAGCAGGCTGGGGACTGAGAGACTCCAGCCCTGTGGGGA 1596
DB      129582 AGGGCTGSCCAGAGGGTGCAGAGCGCGCTGTGTGGGAAGACTCGAGCCCTGGGGGA 129641

QY      1597 TGGTTCGCCAGCTGGAGGGGGTGGGCGCTGGGAGTGGGGAGCCCCCACTTGATCTG 1656
DB      129642 AGTGGCCCAAGGCTGCAGAGGGTGGGCGTGGGCGGGGG-----CTGCTTGCATTG 129694

QY      1657 GTGCCACATTCACTGCAGATCTATGTCGGGCAAGTCAACATGGA-----TG 1702
DB      129695 ATGCCACATTTGCTACAGTCCCAATGTGCAAGTCACTGGAAGATGCCCCCATGGTG 129754

QY      1703 GGGGAAGAAGTTAATAATCTTTGTCAGGAGACCCAGCGACCCCATCAACAACATTGTGAT 1762
DB      129755 GGGCGAGAGTTTCAATAATCTTGTTCAGGAGAGACCACTACCTGTCTATACTTGAAGTG 129814

QY      1763 CTTAGAGGGCGAGAGAGAGGCTGTGAGTGGAGCTGGGAGGCTTTGCCAAGAGTGGCC 1822
DB      129815 CTTTGTCCAGGAGCGG-----GGGAGGCGCTCCCAAGGGCGGCGC 129858

QY      1823 TGTGAGCAGGCGCTCGGAAGATGACAGGCTTTGACAGATGGGAAGTGGGGATGAGAGA 1882
DB      129859 TGTGAG--TGCCCTTGGAAAGCGGCGAGGTTTGTTCAGAGGGGAATGAGGAGGAGAGA 129916

QY      1883 CAGACGAGTGTTCAGGCGCAAGGGAACTGGAAACAAAGAAGAACCT-GAGAATGTAAATCT 1941
DB      129917 CAGGCACAGCATTCAGGCGGAGGAACTGGGGCAAGAAGAATCTCGAGAAATGTAGATCC 129976

QY      1942 ACTTCAAC-CCTGAGCCCTCTTTGCCAAGGGTGCATCTCAGATGCCCTGAAT 1995
DB      129977 ACCTCAACTGCGGACCCCTCACTCGTCCAGGAGTTGCCAACTCCAATGCCCGAAT 130031
```

RESULT 13

AL596246

LOCUS

144136 bp DNA linear

ROD 21-DEC-2002

DEFINITION Mouse DNA sequence from clone RP23-186E14 on chromosome 11, complete sequence.
 ACCESSION AL596246
 VERSION AL596246.10 GI:27369258
 KEYWORDS HTG.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 144136)
 Tracey, A.
 Direct Submission
 Submitted (20-DEC-2002) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
 humquery@sanger.ac.uk
 On Dec 23, 2002 this sequence version replaced gi:18477348.
 During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.
 The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em., EMBL; Sw., SWISSPROT; Tr., TrEMBL; Wp., WORMPEP; Information on the WORMPEP database can be found at http://www.sanger.ac.uk/projects/C_elegans/wormpep RP23-186E14 is from the RPI-23 Mouse PAC Library constructed by the group of Pieter de Jong.
 For further details see <http://www.chori.org/bacpac/home.htm>
 VECTOR: pBACE3.6
 ----- Genome Center
 Center: Wellcome Trust Sanger Institute
 Center code: SC
 Web site: <http://www.sanger.ac.uk>
 Contact: humquery@sanger.ac.uk

 This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC.
 Sequence from the Mouse Genome Sequencing Consortium whole genome shotgun may have been used to confirm this sequence. Sequence data from the whole genome shotgun alone has only been used where it has a phred quality of at least 30.
 Location/Qualifiers
 1..144136
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10090"
 /chromosomes="11"
 /clone="RP23-186E14"
 /clone_lib="RPCL-23"

FEATURES

source
 1..144136
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10090"
 /chromosomes="11"
 /clone="RP23-186E14"
 /clone_lib="RPCL-23"

ORIGIN

Query Match 14.9%; Score 299.8; DB 10; Length 144136;
 Best Local Similarity 65.4%; Pred. No. 1.8e-73;
 Matches 576; Conservative 0; Mismatches 272; Indels 33; Gaps 8;
 QY 774 GTGCTGCGAGTGTGGCCGCGGAGCTCTGTAAAGCCACTGCTGAGAGGCACTC 833
 DB 139203 GTGTGTGCGAGCTGTACCCCGAGG-TGGGCACCTAGGAGCTCATTTGACCGCGTGGCTC 139261
 QY 834 CCATCTTTCTCCATTCTCTAGACTGCTGCTAT-ACAGTCATTTTATGTGTTTC 892
 DB 139262 CCCCATCTCCCTTACCCCGAGAGCTCTGCTGCTGCTTTTGTCAATTCGAAGTGTCTT 139321
 QY 893 GCCAAATTTTATCCAGCTCTGAAATCTCTGAGCTCCC-CTTACAGAGGAGGTGAGCTA 951

RESULT 14
 AC023518/c
 LOCUS AC023518.6
 DEFINITION Mus musculus chromosome 11 clone RP23-321F3 map 11, WORKING DRAFT
 SEQUENCE, 6 unordered pieces.
 ACCESSION AC023518
 VERSION AC023518.6 GI:28913154
 KEYWORDS HTG; HTGS PHASE1; HTGS DRAFT.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 217088)
 AUTHORS Birren, B., Nusbaum, C. and Lander, E.
 TITLE Mus musculus chromosome 11, clone RP23-321F3
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 217088)
 AUTHORS Birren, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N., Anderson, S., Balgwin, J., Barna, N., Beda, F., Boguslavsky, L., Bouckhgalter, B., Brown, A., Burkett, G., Campopiano, A., Castle, A., Choepel, Y., Colangelo, M., Collins, S., Collymore, A., Cooke, P., DeArelano, K., Dewar, K., Dodge, S., Domino, M., Doyle, M.,

Db 139322 ACCAGTTTTATTTCACACTCTGAAACTCAATGACCAAGTTCCCTCAATAGCAGATGAGAG 139381
 QY 952 AGGGCTGGAGCTCAAGGCAATTCAAACCCCTA---CCAGATCTGACGAATGTGATGGCCAC 1008
 Db 139382 TGAACCAAGGCTCAAGCCATCTGACCCCTTCATTCCAGATCTAAACAAATGATGGCCAC 139441
 QY 1009 GTCCCGGAAATATGAAGACCTCTTATGGCATGGGAGGCTGCGAGACAAAGCGGGAG 1068
 Db 139442 ATCCCGGAATATGAAGAAATTCCTATGGCATGGGAGGCTGCGAGACAAAGGTGGGAG 139501
 QY 1069 AGCCATCTCTCAGTTTATCCCGAAATACGTGGAATCTATCAACACAGGCTGCCCGGCTCAA 1128
 Db 139502 AGCCATCTCTCTTTTTCCTCAAGATGATGTGGAGTTCTCCAACAAGATTGCCAAGCTCAA 139561
 QY 1129 TGGTGAAGTCCCTGCTGCGCAACATCATCTGGCACTTGGGTCCC---TTCAATTTTCTCAA 1184
 Db 139562 TGGTGAAGTACCAATGCTGTACCTCTGGCTGGCTGCCAGTACCTCAAACTCTCTCA 139621
 QY 1185 GAGGTGCTGTGAACCCCAAGCCTAGGAAAGGTAGATCCCTGGAGGAGGAGGTA---1240
 Db 139622 GAGGCTCTTGGGATCCCGAGGCCAGGACAAGTAGGGCCCTAGAGGTGTGAGGACATA 139681
 QY 1241 -----ATGTGCTGTGGAGAGCCTGGCTGTGTCTCTCTCTGTAGGCTATGTAGAT 1290
 Db 139682 TTGGGTAAAGTGTGATGGCTGTGGAGAGCTTGACTGTGTCTCTGTGCGAGGCTACACGGAT 139741
 QY 1291 GCAGGGAGCTCTGTGGAGGTCTATGTACGAGACCAATCTCTGGAGCAAGACCTCGAGCGG 1350
 Db 139742 GCAGGGAGTTCATGGAGATCTCTTATACGAGTGTGACCACTGGAGCAAGACCTCGGAAAA 139801
 QY 1351 CTCTTCCAGGAGCTGCAGCCTCTACCTCAACCTGCATGCTAGCTGCGCGGCGCCCTG 1410
 Db 139802 CTGTACAGAGCTGCAGCCTCTACCTGAACCTGTGATGCTCTGTGCTCTCTGCTTCCCTG 139861
 QY 1411 CACCGTCACTACGGGCGCCAGCACATCAACCTGAGGGGCCCACTCTCTCTCACTCTGCTG 1470
 Db 139862 CACCGCCACTATGGTCCGAGTACATCAACCTGGATGGCCCACTCTCTGCCCCATCTGCTA 139921
 QY 1471 GGTAGGGCACTGTGCGGCTTGGAGGGGTAAAGAGCGGACACAGTGTGATGAGGGT 1530
 Db 139922 GGTAGGACATTTGGTCCAGCTTCAGGAGG-----TGGTCAATGGGCAATAAAGGA 139973
 QY 1531 TGGGACAGGCTGACTAGAGGCTAGGAGTACGAGC-AGGCTGGGAGCTGAGAGACTCCAGCCCTG 1589
 Db 139974 CAGTATAGTACTGGCCACAGGCGCAGATCAACCAAGATAGTACAGAACCAAGAGCTC 140033
 QY 1590 TGGGGGATGGTTGCCAGCTGGAGGGGGTGGGCGCTGGG 1630
 Db 140034 TGAGGGAGCGTGGCCCAAGGCTGAGTGGAAAGGGGTGGGG 140074

Fenestor, J., Ferreira, P., FitzHugh, W., Forrest, C., Gage, D.,
 Galagan, J., Gardyna, S., Ginde, S., Goyette, M., Graham, L.,
 Grand-Pierre, N., Grant, G., Hagos, B., Heaford, A., Horton, L.,
 Howland, J. C., Iliev, I., Johnson, R., Jones, C., Kann, L., Karatas, A.,
 Klein, J., Landers, T., Largoczek, K., Lehoczy, J., Levine, R.,
 Lieu, C., Liu, G., Locke, K., Macdonald, P., Marquis, N., McCarthy, M.,
 McEwan, P., McGurk, A., McKernan, K., McPheeters, R., Meldrum, J.,
 Meneus, L., Mihova, T., Miranda, C., Mlenga, V., Morrow, J., Naylor, J.,
 Norman, C. H., O'Connor, T., O'Donnell, P., O'Neill, D., Olivari, T. M.,
 Peterson, K., Pierre, N., Pisani, C., Pollara, V., Raymond, C.,
 Riley, R., Rogov, P., Rothman, D., Roy, A., Santos, R., Schauer, S.,
 Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N.,
 Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J., Tirrell, A.,
 Travers, M., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B.,
 Wu, X., Wyman, D., Ye, W. J., Young, G., Zainoun, J., Zimmer, A. and
 Zody, M.
 Direct Submission
 Submitted (13-FEB-2000) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 3 (bases 1 to 217088)
 Birren, B., Nusbaum, C., Lander, E., Abouelleil, A., Allen, N.,
 Anderson, S., Arachchi, H. M., Barna, N., Bastien, V., Bloom, T.,
 Boguslavskiy, L., Boukhgalter, B., Camarata, J., Chang, J., Choepel, Y.,
 Collymore, A., Cook, A., Cooke, P., Corum, B., DeArelano, K.,
 Diaz, J. S., Dodge, S., Dooley, K., Dorris, L., Erickson, J., Faro, S.,
 Ferreira, P., FitzGerald, M., Gage, D., Galagan, J., Gardyna, S.,
 Graham, L., Grand-Pierre, N., Hafez, N., Hagopian, D., Hagos, B.,
 Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,
 Kanat, A., Karatas, A., Kells, C., Landers, T., Levine, R.,
 Lindblad-Toh, K., Liu, G., Liu, A., Mabbitt, R., Maclean, C.,
 Macdonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M.,
 Meldrum, J., Meneus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J.,
 Nguyen, C., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P.,
 O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N.,
 Rachupka, A., Ramasamy, U., Raymond, C., Retta, R., Rise, C., Rogov, P.,
 Roman, J., Schauer, S., Schnupack, R., Seaman, S., Severy, P., Smith, C.,
 Stange-Thomann, N., Stojanovic, N., Stubbs, M.,
 Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M.,
 Vassiliev, H., Venkataraman, V. S., Viel, R., Vo, A., Wilson, B., Wu, X.,
 Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.
 Direct Submission
 Submitted (11-MAR-2003) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 On Mar 11, 2003 this sequence version replaced gi:11225431.
 All repeats were identified using RepeatMasker:
 Smit, A. F. A. & Green, P. (1996-1997)
 http://ftp.genome.washington.edu/RM/RepeatMasker.html

TITLE JOURNAL COMMENT

Center: Whitehead Institute/ MIT Center for Genome Research
 Center code: WIBR
 Web site: <http://www-seq.wi.mit.edu>
 Contact: sequence_submissions@genome.wi.mit.edu
 ----- Project Information
 Center project name: L3199
 Center clone name: 321_P_3

----- Summary Statistics
 Sequencing vector: M13; M7815; 34% of reads
 Sequencing method: Plasmid; n/a; 66% of reads
 Chemistry: Dye-terminator Big Dye; 100% of reads
 Assembly program: Phrap; version 0.960731
 Consensus quality: 215379 bases at least Q40
 Consensus quality: 215982 bases at least Q30
 Consensus quality: 216295 bases at least Q20
 Insert size: 213000; agarose-fp
 Insert size: 216588; sum-of-contigs
 Quality coverage: 13.3 in Q20 bases; agarose-fp
 Quality coverage: 13.1 in Q20.

* NOTE: This is a 'working draft' sequence. It currently
 * consists of 6 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence

* as soon as it is available and the accession number will
 * be preserved.
 * 1 68732: contig of 68732 bp in length
 * 68733 68832: gap of 100 bp
 * 68833 74687: contig of 5855 bp in length
 * 74688 74787: gap of 100 bp
 * 74788 83104: contig of 8317 bp in length
 * 83105 83204: gap of 100 bp
 * 83205 91107: contig of 7903 bp in length
 * 91108 91207: gap of 100 bp
 * 91208 129624: contig of 38417 bp in length
 * 129625 129724: gap of 100 bp
 * 129725 217088: contig of 87364 bp in length.
 FEATURES
 Location/Qualifiers
 source
 1..217088
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10090"
 /chromosome="11"
 /map="11"
 /clone="RP23-321F3"
 /clone_lib="RPCI-23 Female Mouse BAC"
 1..68732
 /note="assembly_fragment"
 clone_end:SP6
 vector_side:left
 68833..74687
 /note="assembly_fragment"
 74788..83104
 /note="assembly_fragment"
 83205..91107
 /note="assembly_fragment"
 91208..129624
 /note="assembly_fragment"
 129725..217088
 /note="assembly_fragment"
 clone_end:T7
 vector_side:right
 ORIGIN
 Query Match 14.9%; Score 299.8; DB 2; Length 217088;
 Best Local Similarity 65.4%; Pred. No. 1.9e-73; Indels 33; Gaps 8;
 Matches 576; Conservative 0; Mismatches 272;
 QY 774 GTGTCTGAGCATGTGGCCCGGCGGACTCTGTAAAGCACCCTGCTGGAGAGCCTTC 833
 Db 55062 GTGTGTGAGCATGTCAACCCAGG-TGGCACCCCTAGGAGCTCATTCAGCGCGTGC 55004
 QY 834 CCATCTTTTCCCATTTCTCTAGACCTGTGCTTAT-ACAGTCACTTTTATGTGGTTTC 892
 Db 55003 CCCCATTCCCGCTACCCCGCTAGACTTCCTGCGCGTCTTTGTCTATTTCCAGTCTCTT 54944
 QY 893 GCCAATTTTATCCAGCTCTGAATTTCTCTGAGTCCC-CYTAGACAGAGGTGAGCTA 951
 Db 54943 ACCAGTTTTATTCCAACTCTGAACTCAATGACCAAGTTCCCTCAATAGAGATGAGAG 54884
 QY 952 AGGGCTGAGCTCAAGGCATTCAAAACCCCTA---CCAGATCTGACGAATGTATGGCCAC 1008
 Db 54883 TGAACCAAGGCTCAAGCCATCTGACCCCTTCATTCAGATCTAACAAACATGATGGCCAC 54824
 QY 1009 GTCCCGGAAATATGAAGACCTGTGTATGGGATGGGAGGGCTGGCGAGACAGCGGGAG 1068
 Db 54823 ATCCCGGAAATATGAAGAAATTCGTATGGGATGGAAGAGCTGGAGAGCAAGGTGGGAG 54764
 QY 1069 AGCCATCTCCAGTTTTTACCCGAAATACGTGGAATCATCAACAGGCTGCCCGCTCAA 1128
 Db 54763 AGCCATCTCTCTTTTTCCTTTTCCCAAGATGTGGAGTTCTCCACACAGATTGCCAGCTCA 54704
 QY 1129 TGGTGAAGTCCCTGCTGCCAATCATCTGCGACTTTGGGTCCC---TTCTATTTTCTCAAA 1184
 Db 54703 TGGTGAAGTACCAATGCTGGTACCTCTGGCTGGCTGCTCCCGAGTACCTCAAACTCTCA 54644
 QY 1185 GAGTGTGTGTAACCCCAAGCCTAGGAAAAGGTAGATCCCTGGAGGAGGAGGTA--- 1240

Db	54643	
----	-------	--

Mawhney,S., McLeod,M.P., McNeill,T.Z., Meenen,E., Milosavljevic,A., Miner,G.S., Minja,E., Montemayor,J., Moore,S., Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L., Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S., Nwaokemele,O., Okwouu,G., Olarpunsagoon,A., Pal,S., Parks,K., Pasternak,S., Paul,H., Perez,A., Perez,L., Pfanckoch,C., Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.-L., Puazo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R., Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F., Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J., Sanders,W., Savery,G., Scherer,S., Scott,G., Shatsman,S., Shen,H., Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajls,D., Sneed,A., Sodergren,B., Song,X.-Z., Sorelle,R., Taylor,C., Teimle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C., Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K., Valas,R., Vera,V., Villaseana,D., Waldron,L., Walker,B., Wang,J., Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F., Williams,G., Willson,R., Wlezyk,R., Wooden,H., Worley,K., Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V., Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O., Weinstock,G. and Gibbs,R.A.

Direct Submission
Unpublished
2 (bases 1 to 213169)
Worley,K.C.
Direct Submission
Submitted (17-SEP-2001) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 213169)
Rat Genome Sequencing Consortium.
Direct Submission
Submitted (10-MAY-2003) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
On May 10, 2003 this sequence version replaced gi:23321989.
The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GFQE
Center clone name: CH230-44G5
----- Summary Statistics
Assembly program: Atlas 3.0;
Consensus quality: 195455 bases at least Q40
Consensus quality: 199513 bases at least Q30
Consensus quality: 201879 bases at least Q20
Estimated insert size: 207657; sum-of-contigs estimation
Quality coverage: 6x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 3 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence

* as soon as it is available and the accession number will
* be preserved.

1 209047: contig of 209047 bp in length
* 209048 209147: gap of unknown length
* 209148 210639: contig of 1492 bp in length
* 210640 210739: gap of unknown length
* 210740 213169: contig of 2430 bp in length.

FEATURES

source

1..213169
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-44GS"

misc_feature

1..1473
/note="wgs contig"

misc_feature

86723..87837
/note="wgs contig"

misc_feature

143903..145720
/note="wgs contig"

misc_feature

206912..209047
/note="wgs contig"

ORIGIN

Query Match 13.8%; Score 277.2; DB 2; Length 213169;
Best Local Similarity 67.3%; Pred. No. 5e-67;
Matches 479; Conservative 0; Mismatches 203; Indels 30; Gaps 5;
QY 940 AGAGGTGAGCTAGGGCTGAGGCTCAAGGCTTCAAAACCC---CTACCAGATCTGACGAA 996
Db 53506 AGCAGTGAGCAAGAACTGAGGCTCAAAACCTTCGACCCCTTTCTTCCAGATCTGACAAA 53447
QY 997 TGTGATGGCCACGTCCTCCGGAATATGAGACCTGTTATGGCATGGAGGGCTGGCGAGA 1056
Db 53446 TATAATGGCCACGTCCTCCGGAATACGAAGAATTGCTTTGGGTGGAAGAGCTGGCGAGA 53387
QY 1057 CAAGCGGGGAGAGCCATCTCCAGTTTTTACCCGAATACGTGAACTCATCAACACGAGC 1116
Db 53386 CAAGGTGGGAGAGCCATCTTCCCTTTTCCAAAGTAGTGAGACTTCTCCAAAGAT 53327
QY 1117 TGCCCGGCTCAATGGTGGAGTCCCTGTGTCGCAACATCACTGGCACTTGGGTCC---CTTC 1172
Db 53326 CGCCAAAGCTCAACGGTGAGTTACAGTGTGTGTACCTCTGGCTGGCTGCCCCAGCACC 53267
QY 1173 ATTTTCTCAAGAGGTGCTGTCAAAACCCCAAGCCTAGGAAGGTAGATCCCTGGAGGA 1232
Db 53266 TCAAACTCTCAGAGTCTCTCGGATCTCTCAGGCTAGGACAAAGTAGGACCTAGAGGT 53207
QY 1233 GGCAGGTA-----ATGTGTGTTGGGAGAGCCTGGCTGTGTCCCTCTGTGA 1278
Db 53206 GTCAGGGAATATTTGGGTGAGTGACGGTTGTTGGGAGAGCTTGGCTGTCTCTGTGCA 53147
QY 1279 GGCTATGTAGATGACGGGACTCGTGGAGGTCTATGTACGACACCAATCCCTGGAGCAA 1338
Db 53146 GGTACTCTGTATGTCAGGGGATTCCTGGAGATCCTCATATGAGTCCGACCACTTGGAGCAA 53087
QY 1339 GACCTGGAGCGGCTCTTCAGAGAGCTGACGCCACTTCACTCAACCTGGAGGGGCCATTCT 1458
Db 53086 GACCTGGAATAACTATACAGAGAGCTGACGCCCTCTACCTGAACCTGCATGCCTATGTG 53027
QY 1399 CGCCGGGCCCTGCACCGTCACTACGGGGCCAGACATCAACCTGGAGGGGCCATTCT 1458
Db 53026 CGCCGCTCCCTGCACCGCATTTATGGGTCTGAGTACATCAACCTGGATGTTCCCATTCCT 52967
QY 1459 GCTCACCTGCTGGTAAAGGACATGTCGGGCTTGGAGGGGTAAAGACGACACAGT 1518
Db 52966 GCCACCTGCTAGGTAAAGACATCGTCCAGCTTCAAGAGGGTGGACATGGA----- 52914
QY 1519 GTGAGTGGGGTGGGACAGGGCTGACTAGAGGGTAGGGAGC-AGGCTGGGAGCTGAGAG 1577
Db 52913 -TGATAAAGGAACAGATAGTCTGTGCCACAGGGCAGAGATCAAGCCAGAAATAGTGACAG 52855
QY 1578 ACTCCAGCCCTGTGGGGGATGTTGCCAGGCTGGAGGGGGTGGCGCTGG 1629

Db 52854 AACTTGAAGCTCTGAAGGATCATGGCCAAAGCTGGAGTGGAGAGGGGTGAGG 52803

Search completed: November 23, 2004, 04:30:37
Job time : 8338.11 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model
Run on: November 22, 2004, 15:08:20 ; Search time 1233.75 Seconds
(without alignments)
11849.778 Million cell updates/sec

Title: US-10-017-724-1_COPY_3000_5784
Perfect score: 2785
Sequence: 1 gaggacttgacgtgatgg.....aattgtaaaaggtttctt 2785

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4134886 seqs, 2624710521 residues

Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_23Sep04:*

- 1: geneseqn1980s:*
- 2: geneseqn1990s:*
- 3: geneseqn2000s:*
- 4: geneseqn2001as:*
- 5: geneseqn2001bs:*
- 6: geneseqn2002as:*
- 7: geneseqn2002bs:*
- 8: geneseqn2003as:*
- 9: geneseqn2003bs:*
- 10: geneseqn2003cs:*
- 11: geneseqn2003ds:*
- 12: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2785	100.0	5784	6	ABL62660 Colon ade
2	2785	100.0	5784	6	ABZ35081 Human gen
3	2785	100.0	5784	6	ABK10277 Human cDN
4	2785	100.0	5784	6	ABN95662 Gene #216
5	2785	100.0	5784	6	AAD40611 Human thr
6	2785	100.0	5784	6	ABV94768 Human pan
7	2785	100.0	5784	8	ABQ77406 Human THB
8	2785	100.0	5784	9	AAD55115 Human THB
9	2785	100.0	5784	10	ADD31094 Human cDN
10	2785	100.0	5784	10	AD85079 Farnesyl
11	2785	100.0	5784	12	AQ18777 Human sof
12	2745.4	98.6	5843	12	AQ23188 Human sof
13	2735.8	98.2	3787	3	AAC77795 Human can
14	2709.6	97.3	5696	9	ACH03798 Human cDN
15	2709.6	97.3	5696	10	ADJ56219 Bovine cD
16	2456.2	88.2	3522	4	AAH34535 Human col
17	1694	60.8	2780	12	ACF57539 Human thr
18	809.4	29.1	3596	3	AAH88669 Human thr
19	809.4	29.1	3596	10	ABX14061 cDNA enco
20	766.6	27.5	4170	6	ABL35019 Rat cDNA
21	753.6	27.1	4108	2	AAZ32016 Human MET

22	753.6	27.1	4108	5	AAC90073	AAC90073 L07803 cD
23	651.8	23.4	713	4	AAH34534	AAH34534 Human col
c 24	595.6	21.4	619	3	AAZ78040	AAZ78040 CDNA enco
c 25	595.6	21.4	619	4	AAI28778	AAI28778 Colon tum
c 26	595.6	21.4	619	8	ABZ32964	ABZ32964 Human col
c 27	594.8	21.4	848	8	AAZ17196	AAZ17196 Human gen
c 28	540.2	19.4	637	4	AAZ58772	AAZ58772 cDNA #144
c 29	540.2	19.4	540	6	ABV98659	ABV98659 Human pan
c 30	516	18.5	516	6	ABK45682	ABK45682 CDNA enco
c 31	515	18.5	551	6	ABV98694	ABV98694 Human pan
c 32	497.4	17.9	570	6	ABV99051	ABV99051 Human pan
c 33	496.8	17.8	529	6	ABV97469	ABV97469 Human pan
c 34	495.4	17.8	551	6	ABV98796	ABV98796 Human pan
c 35	495.4	17.8	5289	2	AAZ32052	AAZ32052 Human MET
c 36	495.4	17.8	5289	5	AAZ90309	AAZ90309 AB005287
c 37	491.8	17.7	7231	6	ABS62773	ABS62773 Prostate
c 38	491.8	17.7	7231	10	ADJ56211	ADJ56211 Murine cD
c 39	491.8	17.7	7232	6	AAZ94912	AAZ94912 Human DNA
c 40	490.2	17.6	3967	11	ADN39634	ADN39634 Cancer/an
c 41	490.2	17.6	4433	2	AAZ32014	AAZ32014 Human MET
c 42	490.2	17.6	4434	5	AAZ90071	AAZ90071 X04665 CD
c 43	490.2	17.6	4434	10	ADD18445	ADD18445 Human pro
c 44	490.2	17.6	5721	4	AAH02923	AAH02923 Human she
c 45	490.2	17.6	5722	2	AAZ32012	AAZ32012 Human MET

ALIGNMENTS

RESULT 1

ABL62660

ID ABL62660 standard; DNA; 5784 BP.

AC ABL62660;

XX 15-MAY-2002 (first entry)

XX Colon adenocarcinoma related gene sequence SEQ ID NO:997.

XX Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
XX stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
XX cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;
XX gene; ds.

OS Homo sapiens.

XX WO200194629-A2.

PN 13-DEC-2001.

XX 30-MAY-2001; 2001WO-US010838.

XX 05-JUN-2000; 2000US-0209473P.

XX 05-JUN-2000; 2000US-0209531P.

XX 18-SEP-2000; 2000US-0233133P.

XX 18-SEP-2000; 2000US-0233617P.

XX 20-SEP-2000; 2000US-0234009P.

XX 20-SEP-2000; 2000US-0234034P.

XX 22-SEP-2000; 2000US-0234052P.

XX 22-SEP-2000; 2000US-0234509P.

XX 25-SEP-2000; 2000US-0234921P.

XX 25-SEP-2000; 2000US-0234924P.

XX 25-SEP-2000; 2000US-0235077P.

XX 25-SEP-2000; 2000US-0235082P.

XX 25-SEP-2000; 2000US-0235134P.

XX 25-SEP-2000; 2000US-0235280P.

XX 26-SEP-2000; 2000US-0235637P.

XX 26-SEP-2000; 2000US-0235638P.

XX 27-SEP-2000; 2000US-0235711P.

XX 27-SEP-2000; 2000US-0235720P.

XX 27-SEP-2000; 2000US-0235840P.

XX 27-SEP-2000; 2000US-0235863P.

QY	2401	TTGCTGTTGTTTTGCTGCACTTTTACTTTTTTGGGTGTGGAGCTGTATTCCTCGAGACAAC	2460
Db	5400	TTGTGTTGTTTGTCTGCACCTTTTACTTTTTTGGGTGTGGAGCTGTATTCCTCGAGACAAC	5459
QY	2461	GAAGCGTGTGGGATCACTTCATTTAAATGTAGCGACTGTCAAACAGCGTGCAGGTTTTTCTGTTTT	2520
Db	5460	GAAGCGTGTGGGATCACTTCATTTAAATGTAGCGACTGTCAAACAGCGTGCAGGTTTTTCTGTTTT	5519
QY	2521	CTGTGTTGTGGGGTCAACCGGTACAACGTGTGGGAATGACGATGATGTGAATATTTAGAA	2580
Db	5520	CTGTGTTGTGGGGTCAACCGGTACAACGTGTGGGAATGACGATGATGTGAATATTTAGAA	5579
QY	2581	TGTTACCAATATTTTTTGTAAATTAATTTATGTTTTTCTAAACAAATTTATCGTATAGGTGA	2640
Db	5580	TGTTACCAATATTTTTTGTAAATTAATTTATGTTTTTCTAAACAAATTTATCGTATAGGTGA	5639
QY	2641	TGAAACGCTCATGTGTTTTTGGCAAGACTGTAAATATTTATTTATGTGTTCCACATGGTCAA	2700
Db	5640	TGAAACGCTCATGTGTTTTTGGCAAGACTGTAAATATTTATTTATGTGTTCCACATGGTCAA	5699
QY	2701	AATTTACCACTGAAACCCCTGCACCTTAGCTAGTAACCTCATTTTTTAAAGATTAAACAACAGG	2760
Db	5700	AATTTACCACTGAAACCCCTGCACCTTAGCTAGTAACCTCATTTTTTAAAGATTAAACAACAGG	5759
QY	2761	AAATAAATTGTAAAAAGGTTTTTCT 2785	
Db	5760	AAATAAATTGTAAAAAGGTTTTTCT 5784	
RESULT 2			
ABZ35081			
ID	ABZ35081 standard; cDNA; 5784 BP.		
XX	XX		
AC	ABZ35081;		
XX	XX		
XX	XX		
DT	05-FEB-2003 (first entry)		
XX	XX		
DE	Human gene expression profile polynucleotide SEQ ID NO 193.		
XX	XX		
KW	Human; artery; endothelium; umbilical; vein; aorta; pulmonary artery;		
KW	bronchial epithelium; prostate; muscle; lung fibroblast; osteoblast;		
KW	tumour; microarray; genome mapping; antibiotic; antiviral; antifungal;		
KW	gene expression; gene; ss.		
XX	XX		
OS	Homo sapiens.		
XX	XX		
PN	W0200274979-A2.		
XX	XX		
PD	26-SEP-2002.		
XX	XX		
PF	20-MAR-2002; 2002WO-US008456.		
XX	XX		
PR	20-MAR-2001; 2001US-0276947P.		
XX	XX		
PA	(ORTH) ORTHO CLINICAL DIAGNOSTICS INC.		
XX	XX		
PI	Wan J, Wang Y;		
XX	XX		
DR	WPI; 2002-740862/80.		
XX	XX		
PT	New gene expression profile generated from primary, endothelial,		
PT	epithelial, and muscle cell types, useful for identifying disease		
PT	pathologies involving alterations of gene expression, e.g. cancer.		
XX	XX		
PS	Claim 5; Page 401-403; 850pp; English.		
XX	XX		
CC	The invention relates to a gene expression profile comprising one or more		
CC	genes (ABZ34889-ABZ35692) and generated from a cell type. The cell type		
CC	is a coronary artery endothelium, umbilical artery or vein endothelium,		
CC	aortic endothelium, dermal microvascular endothelium, pulmonary artery		
CC	endothelium, myometrium microvascular endothelium, keratinocyte		
CC	epithelium, bronchial epithelium, mammary epithelium, prostate		
CC	epithelium, renal cortical epithelium, renal proximal tubule epithelium.		

4800 TATTTTAAATATGCTAGTAACAACATATGATGATATATTTCTAGAACAAACATCTAATA 4859
1861 AGTATATAATCTCTGCAAAATATGAGCTTGATAATATTAGTTGTGACGATGAACATG 1920
4860 AGTATATAATCTCTGCAAAATATGAGCTTGATAATATTAGTTGTGACGATGAACATG 4919
1921 CTAGAACTGTGAACAGATACATAGAGAAATATGAGGAGTTTATGATGAACCTTAATAT 1980
4920 CTAGAACTGTGAACAGATACATAGAGAAATATGAGGAGTTTATGATGAACCTTAATAT 4979
1981 ATAACTGTCGACGAGATTTAGTTCAATATTTGTTACTGTTATCTATCTGCTGATATGG 2040
4980 ATAACTGTCGACGAGATTTAGTTCAATATTTGTTACTGTTATCTATCTGCTGATATGG 5039
2041 AATCTCTTTAAATCAACGCTGAAACGAATCAGCAATTTAGTCTGCCAGGACACCCAA 2100
5040 AATCTCTTTAAATCAACGCTGAAACGAATCAGCAATTTAGTCTGCCAGGACACCCAA 5099
2101 TAATCAGTCATGTTAATATGACAAAGTTGTTTTGTTTTGTTTTGTTTTGTTTTGTTTTG 2160
5100 TAATCAGTCATGTTAATATGACAAAGTTGTTTTGTTTTGTTTTGTTTTGTTTTGTTTTG 5159
2161 TTTTCTTTGTTTTAAGTTGATCATCTTTCTGCGAGGAATAGTCACTCATCCACTCCACA 2220
5160 TTTTCTTTGTTTTAAGTTGATCATCTTTCTGCGAGGAATAGTCACTCATCCACTCCACA 5219
2221 TAAGGGTTTATGAGAGAGTCTGCTGCTGATGATGATGAGGGGCAAACTCTTTTTC 2280
5220 TAAGGGTTTATGAGAGAGTCTGCTGCTGATGATGATGAGGGGCAAACTCTTTTTC 5279
2281 CCCTTTCTGTTAATAGTCATCATCTTTCTATGCCAAAGGACGATCCATCACTTTAGT 2340
5280 CCCTTTCTGTTAATAGTCATCATCTTTCTATGCCAAAGGACGATCCATCACTTTAGT 5339
2341 CTTAATGATACATGATGATTTGTAATAATTAATTTGTTGTTGTTGTTGTTGTTGTTG 2400
5340 CTTAATGATACATGATGATTTGTAATAATTAATTTGTTGTTGTTGTTGTTGTTGTTG 5399
2401 TTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 2460
5400 TTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 5459
2461 GAAGGTTGGGATATCTCAATTAATGATGCGACTGTCAACGCGTGCAGGTTTCTGTTT 2520
5460 GAAGGTTGGGATATCTCAATTAATGATGCGACTGTCAACGCGTGCAGGTTTCTGTTT 5519
2521 CTGTTCTGTTGGGTCACCGTCAATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 2580
5520 CTGTTCTGTTGGGTCACCGTCAATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 5579
2581 TGTACCATATTTTGTAAATTTATTTATGTTTCTTAAACAAATTTATCGTATAGTTGA 2640
5580 TGTACCATATTTTGTAAATTTATTTATGTTTCTTAAACAAATTTATCGTATAGTTGA 5639
2641 TGAACGTCATGTTGTTGCAAGACATGTAATAATTTATTTATGTTCTCATGTTCA 2700
5640 TGAACGTCATGTTGTTGCAAGACATGTAATAATTTATTTATGTTCTCATGTTCA 5699
2701 AATTTCCACCTGAAACCTCGCATCTAGTACCTCATTTTAAAGATTAAACACAGG 2760
5700 AATTTCCACCTGAAACCTCGCATCTAGTACCTCATTTTAAAGATTAAACACAGG 5759
2761 AAATAAATTTGTAAGGATTTTCT 2785
5760 AAATAAATTTGTAAGGATTTTCT 5784

RESULT 3
ABK10277
ID ABK10277 standard; cDNA; 5784 BP.
XX
AC
XX
XX

DT 21-MAY-2002 (first entry)
XX Human cDNA encoding thrombospondin 2 protein.
DE
XX
XX Human; Osteopontin 1; OPN1; ss; gene; thrombospondin 2; wound response;
KW foreign body response; cut; abrasion; burn; vulnery.
XX
XX Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 240..3758
FT /tag= a
FT /product= "Thrombospondin 2 protein"
XX
XX WO200209735-A2.
XX
XX 07-FEB-2002.
XX
XX 31-JUL-2001; 2001WO-US024147.
XX
XX 01-AUG-2000; 2000US-0222071P.
XX
XX (UNIW) UNIV WASHINGTON.
XX Bornstein P, Kyriakides T, Ratner B, Giachelli C, Martinson L;
XX Scatena M;
XX
XX WPI; 2002-217098/27.
XX P-PSDB; AAU76902.
XX
XX Modulating the amount or biological activity of thrombospondin 2 or
XX osteopontin in an animal for modulating a wound response, comprises
XX introducing osteopontin or thrombospondin 2 antagonist into an animal.
XX
XX Disclosure; Page 50-54; 54pp; English.
XX
XX This invention relates to a method for modulating the amount or
XX biological activity of thrombospondin 2 or osteopontin in an animal. The
XX method involves introducing into the animal an osteopontin or
XX thrombospondin 2 antagonist. Using the methods of the invention the
XX amount or biological activity of thrombospondin 2 or osteopontin protein
XX may be modulated. The method of the invention is useful for modulating
XX the amount or biological activity of thrombospondin 2 or osteopontin in
XX an animal which exhibits a wound response or a foreign body response,
XX where the method can be used to improve the wound response or reduces
XX the foreign body response. The method is useful for modulating a wound
XX response or for reducing a foreign body response in an animal and is also
XX useful for improving the wound response, such as at the site of a cut,
XX abrasion or burn. The present sequence represents the human cDNA encoding
XX the thrombospondin 2 protein used in the method of the invention
XX
SQ Sequence 5784 BP; 1447 A; 1460 C; 1518 G; 1359 T; 0 U; 0 Other;

Query Match 100.0%; Score 2785; DB 6; Length 5784;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2785; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GAGGACTTGGACGGTGTGACGGGGTGATATTTGTAAGATGATTTTGACAATCAAC 60
Db 3000 GAGGACTTGGACGGTGTGACGGGGTGATATTTGTAAGATGATTTTGACAATCAAC 3059
QY 61 ATCCACATATTGATGATGTTGTCCTGTAACCAATGCCATCAGTGACAGACTTCAGG 120
Db 3060 ATCCACATATTGATGATGTTGTCCTGTAACCAATGCCATCAGTGACAGACTTCAGG 3119
QY 121 AACTTCCAGATGGTCCCTTGGATCCCAAGGACCAACCAATTTGATCCCACTGGGTC 180
Db 3120 AACTTCCAGATGGTCCCTTGGATCCCAAGGACCAACCAATTTGATCCCACTGGGTC 3179
QY 181 ATTGCCATCAAGGCAAGGAGCTGGTTTCAGACGACCAACTCGGACCCCGGATCGCTGTA 240
Db 3180 ATTGCCATCAAGGCAAGGAGCTGGTTTCAGACGACCAACTCGGACCCCGGATCGCTGTA 3239

QY 241 GGTGTTGACGAGTTGGCTGTGGACTTCAGTGGACATTTCTACGTAAACACTGACCGG 300
Db |||||
QY 3240 GGTGTTGACGAGTTGGCTGTGGACTTCAGTGGACATTTCTACGTAAACACTGACCGG 3299
Db |||||
QY 301 GACGACGACTATGCTGGCTTGGTCTTTGGTTACAGTCAAGCAGCCGCTTCTATGGGTG 360
Db |||||
QY 3300 GACGACGACTATGCTGGCTTGGTCTTTGGTTACAGTCAAGCAGCCGCTTCTATGGGTG 3359
Db |||||
QY 361 ATGTGGAAGCAGGTGACGACGACTTACTGGAGGACCGCCACGCGGGCTATGGGTAC 420
Db |||||
QY 3360 ATGTGGAAGCAGGTGACGACGACTTACTGGAGGACCGCCACGCGGGCTATGGGTAC 3419
Db |||||
QY 421 TCCGGCGTGTCCCTCAAGGTGGTGAATCCACACAGCGGGACGGGCGACGACCTGAGGAAC 480
Db |||||
QY 3420 TCCGGCGTGTCCCTCAAGGTGGTGAATCCACACAGCGGGACGGGCGACGACCTGAGGAAC 3479
Db |||||
QY 481 GCGTGTGGCACACGGGGAACAACGCCGGGGAGGTGCGAACCTTATGGCAGCACCCGAGG 540
Db |||||
QY 3480 GCGTGTGGCACACGGGGAACAACGCCGGGGAGGTGCGAACCTTATGGCAGCACCCGAGG 3539
Db |||||
QY 541 AACATTGGCTGGAAGGACTACCGGCCCTATAGGTGGCACCTGACTCACAGGCCCCAGAAC 600
Db |||||
QY 3540 AACATTGGCTGGAAGGACTACCGGCCCTATAGGTGGCACCTGACTCACAGGCCCCAGAAC 3599
Db |||||
QY 601 GGCTACATCAGAGTCTTAGTGCATGAAGGAAAAAGGTGATGCGAGACTCAGGACCTATC 660
Db |||||
QY 3600 GGCTACATCAGAGTCTTAGTGCATGAAGGAAAAAGGTGATGCGAGACTCAGGACCTATC 3659
Db |||||
QY 661 TATGACCAAACTACGCTGGCGGGCGGTGGGTCTATTTGCTCTCTCAAGAAATGGTC 720
Db |||||
QY 3660 TATGACCAAACTACGCTGGCGGGCGGTGGGTCTATTTGCTCTCTCAAGAAATGGTC 3719
Db |||||
QY 721 TATTTCTCAGACCTCAAGTACGAATGAGAGATATTTAAACAAGATTTGCTGCAATTCG 780
Db |||||
QY 3720 TATTTCTCAGACCTCAAGTACGAATGAGAGATATTTAAACAAGATTTGCTGCAATTCG 3779
Db |||||
QY 781 GCAATGCCCTGTGATGCTGATGCTGCTCCCTAGACACCTCAGTTCAITGTTGGTCTTTCGGCT 840
Db |||||
QY 3780 GCAATGCCCTGTGATGCTGATGCTGCTCCCTAGACACCTCAGTTCAITGTTGGTCTTTCGGCT 3839
Db |||||
QY 841 TCTCTCTTAGAGACACTCTCTGCTCCCTGACCTTAACTCTGATGGTTCTTCACTCTCTG 900
Db |||||
QY 3840 TCTCTCTTAGAGACACTCTCTGCTCCCTGACCTTAACTCTGATGGTTCTTCACTCTCTG 3899
Db |||||
QY 901 CCAGAACCCCAAAACCCCAAGTGCCTTCAGAGGATAAATCAATCGAACTCAGAGATGAA 960
Db |||||
QY 3900 CCAGAACCCCAAAACCCCAAGTGCCTTCAGAGGATAAATCAATCGAACTCAGAGATGAA 3959
Db |||||
QY 961 CATCTAACCCACTAGAGGAAACCAAGTTTGGTGATATATGAGACTTTATGTGGAGTGA AAA 1020
Db |||||
QY 3960 CATCTAACCCACTAGAGGAAACCAAGTTTGGTGATATATGAGACTTTATGTGGAGTGA AAA 4019
Db |||||
QY 1021 TTGGCGATGCCATTAATGCTGCTTTTCTGTTGTTTAAAGAAATGAGTTTACATATA 1080
Db |||||
QY 4020 TTGGCGATGCCATTAATGCTGCTTTTCTGTTGTTTAAAGAAATGAGTTTACATATA 4079
Db |||||
QY 1081 AAATGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1140
Db |||||
QY 4080 AAATGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4139
Db |||||
QY 1141 CCATTAATGATAAATTAAGCATGAAAAATTTGCTGAACCTACTTTTGGTCTTAAAGTTGT 1200
Db |||||
QY 4140 CCATTAATGATAAATTAAGCATGAAAAATTTGCTGAACCTACTTTTGGTCTTAAAGTTGT 4199
Db |||||
QY 1201 CACTATCTTGAATTAAGTTGCTCTACATGACACACAAATCCCGCTAAATTAATTAATA 1260
Db |||||
QY 4200 CACTATCTTGAATTAAGTTGCTCTACATGACACACAAATCCCGCTAAATTAATTAATA 4259
Db |||||
QY 1261 AACAGGGTCAATTTCAAAATTTGAAGTAATGTTTATAGTAAGGAGAGATTAGAGACAAACAG 1320
Db |||||
QY 4260 AACAGGGTCAATTTCAAAATTTGAAGTAATGTTTATAGTAAGGAGAGATTAGAGACAAACAG 4319
Db |||||
QY 1321 GCATAGCAAAATGACATAAGCTACCGATTAACTAAATCGGAAACATGTAAAAACAGTTACAAAA 1380
Db |||||

Db 4320 GCATAGCAAAATGACATAAGCTACCGATTAACTAAATCGGAAACATGTAAAAACAGTTTACAAA 4379
QY |||||
QY 1381 ATAAACGAACTCTCTCTCTCTCTCAATGAAAGCCCTCATGTGACAGTAGAGATGCAAGTT 1440
Db |||||
QY 4380 ATAAACGAACTCTCTCTCTCTCTCAATGAAAGCCCTCATGTGACAGTAGAGATGCAAGTT 4439
Db |||||
QY 1441 TCATCAAAAGAAACAAACATCTCTGCAATGGGTGTGACGCGGTTCACAGATGTGATTTGGC 1500
Db |||||
QY 4440 TCATCAAAAGAAACAAACATCTCTGCAATGGGTGTGACGCGGTTCACAGATGTGATTTGGC 4499
Db |||||
QY 1501 AAAACCTCATTTAAGTAAAGGTTAGCAGACGAAAGTGCCTGCTTTAGTGTCTGCTTGT 1560
Db |||||
QY 4500 AAAACCTCATTTAAGTAAAGGTTAGCAGACGAAAGTGCCTGCTTTAGTGTCTGCTTGT 4559
Db |||||
QY 1561 GCGTGTGTGCGCTCGGGGAGGCTCTGCTCTGAGCTTCTTCCCGAGCTTTGCTGCTCTGAG 1620
Db |||||
QY 4560 GCGTGTGTGCGCTCGGGGAGGCTCTGCTCTGAGCTTCTTCCCGAGCTTTGCTGCTCTGAG 4619
Db |||||
QY 1621 AGGAACGAGACGACGACGACGAGCGGCGGAAAGGGGCACTAAACCGGTATCTAGGCTTTGG 1680
Db |||||
QY 4620 AGGAACGAGACGACGACGACGAGCGGCGGAAAGGGGCACTAAACCGGTATCTAGGCTTTGG 4679
Db |||||
QY 1681 TAACTGCGGCAAGTCTCTTTTACCTGATTTGATGATACATTTTCAATTAAGGTTCCAGTTA 1740
Db |||||
QY 4680 TAACTGCGGCAAGTCTCTTTTACCTGATTTGATGATACATTTTCAATTAAGGTTCCAGTTA 4739
Db |||||
QY 1741 TAAATATTTTGTAAATATTTTAAAGTGACTATAGAAATGCAAACTCCATTTTACCAGTAAC 1800
Db |||||
QY 4740 TAAATATTTTGTAAATATTTTAAAGTGACTATAGAAATGCAAACTCCATTTTACCAGTAAC 4799
Db |||||
QY 1801 TAAATATTTTGTAAATATTTTAAAGTGACTATAGAAATGCAAACTCCATTTTACCAGTAAC 1860
Db |||||
QY 4800 TAAATATTTTGTAAATATTTTAAAGTGACTATAGAAATGCAAACTCCATTTTACCAGTAAC 4859
Db |||||
QY 1861 AGTATATTAATCTCTGAAATATGAGGCTTCATATATTTAGTTGTACGATGAAGCATG 1920
Db |||||
QY 4860 AGTATATTAATCTCTGAAATATGAGGCTTCATATATTTAGTTGTACGATGAAGCATG 4919
Db |||||
QY 1921 CTAGAAGCTGTAAACAGAAATACATAGAGAAATTAATGAGGATTTTATGATGGAACCTTAATAT 1980
Db |||||
QY 4920 CTAGAAGCTGTAAACAGAAATACATAGAGAAATTAATGAGGATTTTATGATGGAACCTTAATAT 4979
Db |||||
QY 1981 ATAAATGTTGCGAGGATTTTATGTTCAATATTTGTTACTGTTATCTATCTGCTGTATATGG 2040
Db |||||
QY 4980 ATAAATGTTGCGAGGATTTTATGTTCAATATTTGTTACTGTTATCTATCTGCTGTATATGG 5039
Db |||||
QY 2041 AATTTCTTTTAAATCAAAACGCTGAAACGAAATCAGCATTTAGTCTTGCAGGACACACCCAA 2100
Db |||||
QY 5040 AATTTCTTTTAAATCAAAACGCTGAAACGAAATCAGCATTTAGTCTTGCAGGACACACCCAA 5099
Db |||||
QY 2101 TAAATCAGTCAATGTGTAATATGACAAAGTTTGTGTTTGTGTTTGTGTTTGTGTTGTTGG 2160
Db |||||
QY 5100 TAAATCAGTCAATGTGTAATATGACAAAGTTTGTGTTTGTGTTTGTGTTTGTGTTGTTGG 5159
Db |||||
QY 2161 TTTTGTGTTGTTTAAAGTTCATGATCTTCTGACGAAATAGTCACTCATCCACCTCCACA 2220
Db |||||
QY 5160 TTTTGTGTTGTTTAAAGTTCATGATCTTCTGACGAAATAGTCACTCATCCACCTCCACA 5219
Db |||||
QY 2221 TAAAGGTTTAAAGTTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2280
Db |||||
QY 5220 TAAAGGTTTAAAGTTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5279
Db |||||
QY 2281 CCGTCTTCTGTTTAAATAGTTCATCATTTTCTATGCCAAAACAGGAAACGATCCATTAACCTTTAGT 2340
Db |||||
QY 5280 CCGTCTTCTGTTTAAATAGTTCATCATTTTCTATGCCAAAACAGGAAACGATCCATTAACCTTTAGT 5339
Db |||||
QY 2341 CTTAATGTACACATTTGCAATTTTGAATAAATTAATTTTGTGTTTGTGTTTGTGTTTGTGTTG 2400
Db |||||
QY 5340 CTTAATGTACACATTTGCAATTTTGAATAAATTAATTTTGTGTTTGTGTTTGTGTTTGTGTTG 5399
Db |||||
QY 2401 TTGTGTTGTTTGTGTTGTTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTT 2460
Db |||||

Db 5400 TTGTTGTTGTTGCTGCACTTTTACTTTTTTTCGTTGGAGCTGTAATCCCGAGACAAC 5459
QY 2461 GAAGCGTTGGGATCTTCAATTAATGTAGCACTGTCAACAGCGTCAGGTTTCTGTTT 2520
Db 5460 GAAGCGTTGGGATCTTCAATTAATGTAGCACTGTCAACAGCGTCAGGTTTCTGTTT 5519
QY 2521 CTGTTGTTGGGCTCAACCGTACAATGTTGTTGGGAATGACGATGATGTAATATTAGAA 2580
Db 5520 CTGTTGTTGGGCTCAACCGTACAATGTTGTTGGGAATGACGATGATGTAATATTAGAA 5579
QY 2581 TGTACCATATTTTGTAAATATTATTTATGTTTCTAAACAATTTATCGTATAGTTGA 2640
Db 5580 TGTACCATATTTTGTAAATATTATTTATGTTTCTAAACAATTTATCGTATAGTTGA 5639
QY 2641 TGAACGTCATGTTGTTTGCCTAAAGACTGTAAATATTATTTATGTTTCAATGTCAA 2700
Db 5640 TGAACGTCATGTTGTTTGCCTAAAGACTGTAAATATTATTTATGTTTCAATGTCAA 5699
QY 2701 AATTTCAACCTGAAACCTGCACTTAGCTAGAACCTCATTTTAAAGATTAAACAACAGG 2760
Db 5700 AATTTCAACCTGAAACCTGCACTTAGCTAGAACCTCATTTTAAAGATTAAACAACAGG 5759
QY 2761 AATTAATTTGTAATAAGGTTTCT 2785
Db 5760 AATTAATTTGTAATAAGGTTTCT 5784

RESULT 4

ABN95662
ID ABN95662 standard; DNA; 5784 BP.
XX AC ABN95662;
XX DT 13-AUG-2002 (first entry)
XX DE Gene #2160 used to diagnose liver cancer.
XX KW Gene; liver cancer; ds; hepatocellular carcinoma; hepatotropic;
KW metastatic liver tumour; cytostatic; expression profile; disease state;
KW disease progression; drug toxicity; drug efficacy; drug metabolism.
XX OS Homo sapiens.
XX PN WO200229103-A2.
XX PD 11-APR-2002.
XX PF 02-OCT-2001; 2001WO-US030589.
XX PR 02-OCT-2000; 2000US-0237054P.
XX PA (GENE-) GENE LOGIC INC.
XX PI Horne D, Alvares C, Peres-Da-Silva S, Vockley JG;
XX WPI; 2002-426119/45.

Diagnosing and detecting the progression of liver cancer, hepatocellular carcinoma or metastatic liver tumor in a patient, involves detecting the level of expression of two or more genes in a liver tissue sample.

Claim 1; SEQ ID NO 2160; 298pp; English.

The invention relates to a novel method for diagnosing and detecting the progression of liver cancer, hepatocellular carcinoma or metastatic liver tumour in a patient, and differentiating metastatic liver cancer from hepatocellular carcinoma in a patient, involving detecting the level of expression of two or more genes represented in ABN93503-ABN97455 in a tissue sample. The method of the invention has hepatotropic, and cytostatic activity. The method is useful for diagnosing and detecting the progression of liver cancer, hepatocellular carcinoma and metastatic liver carcinoma in a patient. The method is useful for identifying expression profiles which serve as useful diagnostic markers as well as

CC markers that can be used to monitor disease states, disease progression, drug toxicity, drug efficacy and drug metabolism. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

SQ Sequence 5784 BP; 1447 A; 1460 C; 1518 G; 1359 T; 0 U; 0 Other;

Query Match 100.0%; Score 2785; DB 6; Length 5784;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 2785; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAGGACTTCGACGCTGATGCGGCGGTGATATTGTAAGATGATTTGACATGACAAC 60
Db 3000 GAGGACTTCGACGCGTGATGACGCGGTGATATTGTAAGATGATTTGACATGACAAC 3059
QY 61 ATCCAGATATTGATGATGTTGTCCTGAAAAAATGCCATCAGTGAGACAGACTTCAGG 120
Db 3060 ATCCAGATATTGATGATGTTGTCCTGAAAAAATGCCATCAGTGAGACAGACTTCAGG 3119
QY 121 AACTTCCAGATGTCCTCCCTTGGATCCCAAGGACACCCAAATTTGATCCCAACTGGGTC 180
Db 3120 AACTTCCAGATGTCCTCCCTTGGATCCCAAGGACACCCAAATTTGATCCCAACTGGGTC 3179
QY 181 ATTCCGCAATCAAGGCAAGGAGCTGTTTCAGACAGCCAACTCGGACCCCGGCAATCGCTGTA 240
Db 3180 ATTCCGCAATCAAGGCAAGGAGCTGTTTCAGACAGCCAACTCGGACCCCGGCAATCGCTGTA 3239
QY 241 GGTTTTGACAGTTTGGGTCTGTGACTTCAGTGGCACATTTAGCTAAACACTGACCGG 300
Db 3240 GGTTTTGACAGTTTGGGTCTGTGACTTCAGTGGCACATTTAGCTAAACACTGACCGG 3299
QY 301 GACGACGATATGCTGGCTTCTGTTGGTTACGAGTCAAGCAGCGCTTCTATGTGGTG 360
Db 3300 GACGACGATATGCTGGCTTCTGTTGGTTACGAGTCAAGCAGCGCTTCTATGTGGTG 3359
QY 361 ATGTGGAAGCAGGTGACGACAGCTACTGGGAGGACAGCCACCGGGGCTTATGGCTAC 420
Db 3360 ATGTGGAAGCAGGTGACGACAGCTACTGGGAGGACAGCCACCGGGGCTTATGGCTAC 3419
QY 421 TCCGCGCTGTCCCTCAAGGTGGTGAATCCACCGGGGACGGGCGAGCACCTGAGGAC 480
Db 3420 TCCGCGCTGTCCCTCAAGGTGGTGAATCCACCGGGGACGGGCGAGCACCTGAGGAC 3479
QY 481 GCGCTGTGGGACACGGGGAACAACCGCGGGGAGGTGCGAACTTATGGACGACCCACAGG 540
Db 3480 GCGCTGTGGGACACGGGGAACAACCGCGGGGAGGTGCGAACTTATGGACGACCCACAGG 3539
QY 541 AACATTGGCTGGAAGGACTACACGCGCTTATAGTGGGACCTGACTCAGGCGCCCAAGACC 600
Db 3540 AACATTGGCTGGAAGGACTACACGCGCTTATAGTGGGACCTGACTCAGGCGCCCAAGACC 3599
QY 601 GGCTACATCAGAGTCTTAGTGCATGAAGAAAAACAGGTGATGGGACGACTCAGGACCTATC 660
Db 3600 GGCTACATCAGAGTCTTAGTGCATGAAGAAAAACAGGTGATGGGACGACTCAGGACCTATC 3659
QY 661 TATGACCAAACTCAGCTGGCGGCGCTGGTCTATTTGTTCTTCTCAAGAAATGGTC 720
Db 3660 TATGACCAAACTCAGCTGGCGGCGCTGGTCTATTTGTTCTTCTCAAGAAATGGTC 3719
QY 721 TATTTCTCAGACCTCAAGTACGATGAGATATTTAAACAAGATTGCTGCAATTTCCG 780
Db 3720 TATTTCTCAGACCTCAAGTACGATGAGATATTTAAACAAGATTGCTGCAATTTCCG 3779
QY 781 GCAATGCGCTGTGATGCGCATGGTCCCTAGACACTCAGTTTCAATTTGGTCTTGGGCT 840
Db 3780 GCAATGCGCTGTGATGCGCATGGTCCCTAGACACTCAGTTTCAATTTGGTCTTGGGCT 3839
QY 841 TCTCTCTCAGACGACTCTCTGTCCTTGAACCTTAACTCTGATGGTTTCTTCACTCTG 900
Db 3840 TCTCTCTCAGACGACTCTCTGTCCTTGAACCTTAACTCTGATGGTTTCTTCACTCTG 3899
QY 901 CCAGCAACCCCAACCCCAAGTGCCTTCAGAGGATAAATATCAATGGAACCTCAGAGATGAA 960

D _b	3900	CCAGCAACCCCAAAACCCAAAGTCCTTCAGAGGATAAATATCAATCGAACTCAGAGATGA	3959
Q _y	961	CATCTAACCCCATAGAGGAAAAACAGTTTGGTGATATATAGAGACTTTATGTGGAGTGA	1020
D _b	3960	CATCTAACCCCATAGAGGAAAAACAGTTTGGTGATATATAGAGACTTTATGTGGAGTGA	4019
Q _y	1021	TTGGGCATGCCATTACATTCGCTTTTTCTGTGTTGTTTAAAAAGAATGAGCTTTACATATA	1080
D _b	4020	TTGGGCATGCCATTACATTCGCTTTTTCTGTGTTGTTTAAAAAGAATGAGCTTTACATATA	4079
Q _y	1081	AAATGTAATTACTTATTTATTTATGTGTATATGGAGTTGAAGGGAATACTGTGCATAAG	1140
D _b	4080	AAATGTAATTACTTATTTATTTATGTGTATATGGAGTTGAAGGGAATACTGTGCATAAG	4139
Q _y	1141	CCATTATGATAAATTAAGCATGAAAAATATTGCTGAACTACTTTTGGTGCTTTAAAGTTGT	1200
D _b	4140	CCATTATGATAAATTAAGCATGAAAAATATTGCTGAACTACTTTTGGTGCTTTAAAGTTGT	4199
Q _y	1201	CACATTCTTGAATTAGAGTTGCTCPACAATGACACACAAATCCCGCTAAAAATAATTATA	1260
D _b	4200	CACATTCTTGAATTAGAGTTGCTCPACAATGACACACAAATCCCGCTAAAAATAATTATA	4259
Q _y	1261	AACAAGGGTCAATTCAAATTTGAAGTAAATGTTTTAGTAAGGAGAGATTAGAACACAACAG	1320
D _b	4260	AACAAGGGTCAATTCAAATTTGAAGTAAATGTTTTAGTAAGGAGAGATTAGAACACAACAG	4319
Q _y	1321	GCATAGCAAATGACATAAGCTACCGATTAACTAATCGGAACATGTHAAAACAGTTACAAA	1380
D _b	4320	GCATAGCAAATGACATAAGCTACCGATTAACTAATCGGAACATGTHAAAACAGTTACAAA	4379
Q _y	1381	ATAAACGAATCTCTCTTGTCTCTACAACTGAAGCCCTCATGTGCAGTAGAGATGCAGTT	1440
D _b	4380	ATAAACGAATCTCTCTTGTCTCTACAACTGAAGCCCTCATGTGCAGTAGAGATGCAGTT	4439
Q _y	1441	TCATCAAGAAACAAACATCTCTGCAAAATGGGTGTGACGGGTTCACAGATGTGGATTGGC	1500
D _b	4440	TCATCAAGAAACAAACATCTCTGCAAAATGGGTGTGACGGGTTCACAGATGTGGATTGGC	4499
Q _y	1501	AAAACTCATTTAAGTAAAAAGTTAGCAGAGCAAAAGTGGGTGCTTTAGCTGCTGTTGT	1560
D _b	4500	AAAACTCATTTAAGTAAAAAGTTAGCAGAGCAAAAGTGGGTGCTTTAGCTGCTGTTGT	4559
Q _y	1561	GCGGTGTGGGTGCGGGAGGCTCCTGCCTGAGCTCTTCCCCAGCTTTGCTGCCCTGAG	1620
D _b	4560	GCGGTGTGGGTGCGGGAGGCTCCTGCCTGAGCTCTTCCCCAGCTTTGCTGCCCTGAG	4619
Q _y	1621	AGGAACACAGACGACAGCACAGGCGGMAAGGCGCATCTAAGCGGTATCTAGGCTTTGG	1680
D _b	4620	AGGAACACAGACGACAGCACAGGCGGMAAGGCGCATCTAAGCGGTATCTAGGCTTTGG	4679
Q _y	1681	TAACTGCGGCAAGATTGCTTTTAACTGATTTGATGATACATTTCAATTAAGTTCCAGTTA	1740
D _b	4680	TAACTGCGGCAAGATTGCTTTTAACTGATTTGATGATACATTTCAATTAAGTTCCAGTTA	4739
Q _y	1741	TAAATATTTTGTAAATTTATTTAAGTGACATATAGATGCAACTCCCAATTTACCAGTAACT	1800
D _b	4740	TAAATATTTTGTAAATTTATTTAAGTGACATATAGATGCAACTCCCAATTTACCAGTAACT	4799
Q _y	1801	TATTTTAAATATGCTTAGTACAACATATGTAGTATAATTTCTAGAAACAAACATCTAATA	1860
D _b	4800	TATTTTAAATATGCTTAGTACAACATATGTAGTATAATTTCTAGAAACAAACATCTAATA	4859
Q _y	1861	AGTATATAATCCTGTGAAAAATATGAGGCTTGATATAATTAGTTGTGCACATGAGCATG	1920
D _b	4860	AGTATATAATCCTGTGAAAAATATGAGGCTTGATATAATTAGTTGTGCACATGAGCATG	4919
Q _y	1921	CTAGAGCTCTTAACAGATATACATAGAGAAATATCAGGAGTTTATGATGGAACTTAATAT	1980
D _b	4920	CTAGAGCTCTTAACAGAGATAACATAGAGAAATATCAGGAGTTTATGATGGAACTTAATAT	4979
Q _y	1981	ATAATGTTGCCAGGCCATTTTGTGTTCAATATTTGTTTACTGTTATCTATCTGCTGTATAGG	2040

Db	4980	ATAAGTGTGCCAGAGATTTTAGTTCOAATATTGTGTACTGTTATCTATCTGCTGTATAGG	5039
QY	2041	AAATCTTTTAAATTCAAAACGCTGAAACGAATCAGCAATTTAGTCTTTGCGAGGCACACCCAA	2100
Db	5040	AAATCTTTTAAATTCAAAACGCTGAAACGAATCAGCAATTTAGTCTTTGCGAGGCACACCCAA	5099
QY	2101	TAAACAGTCATGTGTAAATGATGACAAGTTTGTGTTTTGTTTTGTTTTTGTGTTGTTGGTTGG	2160
Db	5100	TAAACAGTCATGTGTAAATGATGACAAGTTTGTGTTTTGTTTTGTTTTTGTGTTGTTGGTTGG	5159
QY	2161	TTTCTTTTGTCTTAAAGTGTGCATGATCTTCTGCGAGAAATAGTCATCTCATCCACATCCACA	2220
Db	5160	TTTCTTTTGTCTTAAAGTGTGCATGATCTTCTGCGAGAAATAGTCATCTCATCCACATCCACA	5219
QY	2221	TAAAGGGTTTGTAGTAAGAGAAAGTCTGTCGTCTGTATGATGGATAGGGGCAAAATCTTTTTC	2280
Db	5220	TAAAGGGTTTGTAGTAAGAGAAAGTCTGTCGTCTGTATGATGGATAGGGGCAAAATCTTTTTC	5279
QY	2281	CCCTTTCTGTTAAATGATCATCATTTCTATGCCAAAACAGGAACGATCCATAACTTTAGT	2340
Db	5280	CCCTTTCTGTTAAATGATCATCATTTCTATGCCAAAACAGGAACGATCCATAACTTTAGT	5339
QY	2341	CTTAATGTACACATTCGATTTTGATAAAATAAATTTGTTGTTCTCTTTGAGTTGATCG	2400
Db	5340	CTTAATGTACACATTCGATTTTGATAAAATAAATTTGTTGTTCTCTTTGAGTTGATCG	5399
QY	2401	TTGTGTTGTTTTGCTGCACCTTTTTTACTTTTTTTCGCTGTGGAGCTGTATTTCCGAGACAAC	2460
Db	5400	TTGTGTTGTTTTGCTGCACCTTTTTTACTTTTTTTCGCTGTGGAGCTGTATTTCCGAGACAAC	5459
QY	2461	GAAGCGTTGGGATPACTTCAATTAATGTAGCGACTGTCAACAGCGTGCAGGTTTTCTGTTT	2520
Db	5460	GAAGCGTTGGGATPACTTCAATTAATGTAGCGACTGTCAACAGCGTGCAGGTTTTCTGTTT	5519
QY	2521	CTGTGTTGTTGGGTCAACCGTACAAATGGTGTGGGAATGACGATGATGTGAATATTAGAA	2580
Db	5520	CTGTGTTGTTGGGTCAACCGTACAAATGGTGTGGGAATGACGATGATGTGAATATTAGAA	5579
QY	2581	TGTACCATATTTTTTGTAAATTAATTTATGTTTTCTTCAACAAATTTATCGTAGTTCGA	2640
Db	5580	TGTACCATATTTTTTGTAAATTAATTTATGTTTTCTCAACAAATTTATCGTAGTTCGA	5639
QY	2641	TGAAACGTCATGTGTTTTTGCCAAAGACTGTGAATAATTTATTTATGTTTCATGTGTCAA	2700
Db	5640	TGAAACGTCATGTGTTTTTGCCAAAGACTGTGAATAATTTATTTATGTTTCATGTGTCAA	5699
QY	2701	AAATTCACCACTGAACCCCTGCACCTTAGCTAGAACCTCATTTTTTAAAGATTAAACACAG	2760
Db	5700	AAATTCACCACTGAACCCCTGCACCTTAGCTAGAACCTCATTTTTTAAAGATTAAACACAG	5759
QY	2761	AAATAAATTCGAAAAAGGTTTTCT	2785
Db	5760	AAATAAATTCGAAAAAGGTTTTCT	5784
RESULT 5			
AAAD40611			
ID	AAAD40611	standard; DNA; 5784 BP.	
XX	AC	AAAD40611;	
XX	DT	30-OCT-2002 (first entry)	
XX	DE	Human thrombospondin (TSP)-2 DNA.	
XX	KW	Human; cardiovascular disease; CVD; diagnosis; thrombospondin; TSP;	
XX	KW	anticoagulant therapy; atherosclerosis; coronary artery disease; MI;	
XX	KW	myocardial infarction; venous thromboembolism; pulmonary embolism;	
XX	KW	peripheral vascular disease; congestive heart failure; gene therapy;	
XX	KW	stroke; TSP-2; single nucleotide polymorphism; SNP; gene; ds.	
OS	XX	Homo sapiens.	

Key Location/Qualifiers
CDS 240..3758
variation /tag= a
/product= "TSP-2 protein"
replace(3949, G)
/tag= b
/standard_name= "Single nucleotide polymorphism (SNP)"
WO200239122-A2.
16-MAY-2002.
09-NOV-2001; 2001WO-US047406.
09-NOV-2000; 2000US-0248185P.
22-DEC-2000; 2000US-0257417P.
(MILL-) MILLENNIUM PHARM INC.
McCarthy J;
WPI; 2002-519324/55.
P-PSDB; AAB25031.
Diagnosing and monitoring a cardiovascular disease in a patient, by
assessing the level of thrombospondin protein and/or gene expression in a
biological sample.
Example 1; Page 100-102; 108pp; English.
The invention relates to a method of diagnosing and monitoring a
cardiovascular disease (CVD) in a patient. The method involves assessing
the level of thrombospondin (TSP) protein and/or gene expression in a
biological sample. The invention also provides a method useful for
assessing the efficacy of a compound or anticoagulant therapy for
inhibiting CVD in a patient. CVD include atherosclerosis, coronary artery
disease, myocardial infarction (MI), stroke, peripheral vascular disease,
congestive heart failure, venous thromboembolism and pulmonary embolism.
The invention is also used in gene therapy. The present sequence is human
TSP-2 DNA
Sequence 5784 BP; 1447 A; 1460 C; 1518 G; 1359 T; 0 U; 0 Other;
Query Match 100.0%; Score 2785; DB 6; Length 5784;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2785; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 GAGGACTGGACGGTGATGGACGGGGTGATATTTGTAAGATGATTTTGACAATGACAAC 60
3000 GAGGACTGGACGGTGATGGACGGGGTGATATTTGTAAGATGATTTTGACAATGACAAC 3059
61 ATCCGAGATATTGATGATGTGTGCTGCTGAAACCAATGCCATCAGTGACAGACTTCAGG 120
3060 ATCCGAGATATTGATGATGTGTGCTGCTGAAACCAATGCCATCAGTGACAGACTTCAGG 3119
121 AACTTCAGATGGTCCCTTGGATGCCAAGGGACCAACCAATTTGATCCCACTGGGTC 180
3120 AACTTCCAGATGGTCCCTTGGATGCCAAGGGACCAACCAATTTGATCCCACTGGGTC 3179
181 ATTCGCCATCAAGCAGGAGCTGGTTCAGACGCCAATCTGGACCCCGGATCGCTGTA 240
3180 ATTCGCCATCAAGCAGGAGCTGGTTCAGACGCCAATCTGGACCCCGGATCGCTGTA 3239
241 GGTTTTACGAGTTGGTCTGTGGATCTCAGTGCGACATTTCTACGTAACACTGACCGG 300
3240 GGTTTTACGAGTTGGTCTGTGGATCTCAGTGCGACATTTCTACGTAACACTGACCGG 3299
301 GACGACACTATGCTGGCTGTCTTTGGTTACAGTCAAGCAGCCGCTTCTATGTGGTG 360
3300 GACGACACTATGCTGGCTGTCTTTGGTTACAGTCAAGCAGCCGCTTCTATGTGGTG 3359
361 ATCTGGAGAGGTGACGACGACCTACTCTGGAGGACCGAGCCCGGCTTATGGCTAC 420

Db 3360 ATCTGGAACGAGGTGACGACGACCTACTGGAGGACCGCCACCGGGGCTATGGCTAC 3419
QY 421 TCCGGGCTGTCCCTCAAGGTGGTGAATCCACACCGGGGACCGGGCGAGCACCTTGAGGAAC 480
Db 3420 TCCGGGCTGTCCCTCAAGGTGGTGAATCCACACCGGGGACCGGGCGAGCACCTTGAGGAAC 3479
QY 481 GGGCTGTGGCACAACCGGGGACACCGGGGCGAGGTGCGAACCTTATGGCAGACCCAGG 540
Db 3480 GGGCTGTGGCACAACCGGGGACACCGGGGCGAGGTGCGAACCTTATGGCAGACCCAGG 3539
QY 541 AACATTGGCTGAAGGACTACACGGCTATAGTGGCACCCTGACTCAGAGGCCAGAGACC 600
Db 3540 AACATTGGCTGAAGGACTACACGGCTATAGTGGCACCCTGACTCAGAGGCCAGAGACC 3599
QY 601 GGCTACATCAGAGTCTTATGTGATGAAGGAAACAGGTTCATGGCAGACTCAGGACCTATC 660
Db 3600 GGCTACATCAGAGTCTTATGTGATGAAGGAAACAGGTTCATGGCAGACTCAGGACCTATC 3659
QY 661 TATGACCAAACTACGCTGGGGGCGGCTGGGTCTATTGTCTTCTCAAGAAATGGTC 720
Db 3660 TATGACCAAACTACGCTGGGGGCGGCTGGGTCTATTGTCTTCTCAAGAAATGGTC 3719
QY 721 TATTTCTCAGACCTCAAGTACGAATGCAGAGATATTTAAACAAGATTTTGTGCTTCCG 780
Db 3720 TATTTCTCAGACCTCAAGTACGAATGCAGAGATATTTAAACAAGATTTTGTGCTTCCG 3779
QY 781 GCAATGCCCTGTGCATGCCATGGTCCCTAGACACCTCAGTTCAITTTGGTCTTCCGCGCT 840
Db 3780 GCAATGCCCTGTGCATGCCATGGTCCCTAGACACCTCAGTTCAITTTGGTCTTCCGCGCT 3839
QY 841 TCTCTCTCAGACGACCTCTCTCTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTG 900
Db 3840 TCTCTCTCAGACGACCTCTCTCTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTG 3899
QY 901 CCAGCAACCCCAACCCCAAGTGCCTTCAGAGATATAATCAATGGAATCAGAGATGAA 960
Db 3900 CCAGCAACCCCAACCCCAAGTGCCTTCAGAGATATAATCAATGGAATCAGAGATGAA 3959
QY 961 CATCTAACCCCTAGAGGAAACCAAGTTGGTGATATATGAGACTTTATGTGGAGTGAATA 1020
Db 3960 CATCTAACCCCTAGAGGAAACCAAGTTGGTGATATATGAGACTTTATGTGGAGTGAATA 4019
QY 1021 TTGGGCAATGCAATTCATTTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTG 1080
Db 4020 TTGGGCAATGCAATTCATTTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTG 4079
QY 1081 AAATGTAAATTAATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 1140
Db 4080 AAATGTAAATTAATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 4139
QY 1141 CCATTATGATAAATTAAGCATGAAAAATAATTTGCTGAACTACTTTTGGTGTCTTAAAGTTGT 1200
Db 4140 CCATTATGATAAATTAAGCATGAAAAATAATTTGCTGAACTACTTTTGGTGTCTTAAAGTTGT 4199
QY 1201 CACTATTCTTGAATAGAGTGTCTTCAATATGACACAAATCCCGCTAAATAAATATA 1260
Db 4200 CACTATTCTTGAATAGAGTGTCTTCAATATGACACAAATCCCGCTAAATAAATATA 4259
QY 1261 AACAGGGTCAATTCAAATTTGAAGTAATGTTTATGTAAGGAGAGATTAGAGACACAG 1320
Db 4260 AACAGGGTCAATTCAAATTTGAAGTAATGTTTATGTAAGGAGAGATTAGAGACACAG 4319
QY 1321 GCATAGCAATGACATAAGCTACCGATTAACTAATCGAACTGTAAAAACAGTTTACAAAA 1380
Db 4320 GCATAGCAATGACATAAGCTACCGATTAACTAATCGAACTGTAAAAACAGTTTACAAAA 4379
QY 1381 ATAAACGACTCTCTCTTGTCTTCAATGAAAGCCCTCATGTGCGAGTAGAGATGCAATT 1440
Db 4380 ATAAACGACTCTCTCTTGTCTTCAATGAAAGCCCTCATGTGCGAGTAGAGATGCAATT 4439
QY 1441 TCATCAAGAAACAAACATCTTGTGCAATGGGTGTGACGGGTTCCAGATGTGATTTGGC 1500
Db 4440 TCATCAAGAAACAAACATCTTGTGCAATGGGTGTGACGGGTTCCAGATGTGATTTGGC 4499

QY	1501	AAAACCTCATTTTAAAGTAAAGGTTAGCAGACAAAGTGGCGTCTCTTTAGTCTGCTGCTGT	1560	QY	2581	TGTACCATATTTTGTGTAATTTATTTATCTTTTCTAAACAAATTTATCGTATAGGTTGA	2640
Db	4500	AAAACCTCATTTTAAAGTAAAGGTTAGCAGACAAAGTGGCGTCTCTTTAGTCTGCTGCTGT	4559	Db	5580	TGTACCATATTTTGTGTAATTTATTTATCTTTTCTAAACAAATTTATCGTATAGGTTGA	5639
QY	1561	GCGTGTGTGGCGTGGGGAGGCTCCCTGCGCTAGCTTCCTTCCCGAGCTTTGTCGCTGAG	1620	QY	2641	TGAAACGTCATGTGTTTGTCCAAAGACTGTAAATATTTTATTTATGTTGTTTCAATGTC	2700
Db	4560	GCGTGTGTGGCGTGGGGAGGCTCCCTGCGCTAGCTTCCTTCCCGAGCTTTGTCGCTGAG	4619	Db	5640	TGAAACGTCATGTGTTTGTCCAAAGACTGTAAATATTTTATTTATGTTGTTTCAATGTC	5699
QY	1621	AGGAACAGAGCAGACGACAGCGCGGAAAGGCGGATCTAAACGCGTATCTAGGCTTTGG	1680	QY	2701	AAATTTCCACCACTGAAACCCCTGCACTTTAGCTAGAACCTCATTTTAAAGATTAAACACAG	2760
Db	4620	AGGAACAGAGCAGACGACAGCGCGGAAAGGCGGATCTAAACGCGTATCTAGGCTTTGG	4679	Db	5700	AAATTTCCACCACTGAAACCCCTGCACTTTAGCTAGAACCTCATTTTAAAGATTAAACACAG	5759
QY	1681	TAACTGGGACAGTGTCTTTTACCTGATTTGATGATACATTTTCAATTAAGGTTCCAGTTA	1740	QY	2761	AAATAAATTTGTAATAAAGGTTTTCT	2785
Db	4680	TAACTGGGACAGTGTCTTTTACCTGATTTGATGATGATACATTTTCAATTAAGGTTCCAGTTA	4739	Db	5760	AAATAAATTTGTAATAAAGGTTTTCT	5784
QY	1741	TAAATATTTTGTAAATATTTTAAAGTGACTATAGAAATCAACTCCATTTACCAAGTAACT	1800	RESULT 6			
Db	4740	TAAATATTTTGTAAATATTTTAAAGTGACTATAGAAATCAACTCCATTTACCAAGTAACT	4799	ABV94768			
QY	1801	TATTTTAAATGCTAGTAAACATATGATATGATATTTCTAGAAACAAACATCTAATA	1860	ID	ABV94768	standard; cDNA; 5784 BP.	
Db	4800	TATTTTAAATGCTAGTAAACATATGATATGATATTTCTAGAAACAAACATCTAATA	4859	XX	ABV94768;		
QY	1861	AGTATATATCTCTGAAATATGAGCTTGATATATTTAGTGTGTCCAGATGAAGCATG	1920	AC	ABV94768;		
Db	4860	AGTATATATCTCTGAAATATGAGCTTGATATATTTAGTGTGTCCAGATGAAGCATG	4919	XX	14-JAN-2003 (first entry)		
QY	1921	CTAGAAGCTGTAACAGAAATACATAGAAATATGAGGAGTTTATGATGAACCTTAATAT	1980	DT	Human pancreatic cancer expressed cDNA SEQ ID NO 149.		
Db	4920	CTAGAAGCTGTAACAGAAATACATAGAAATATGAGGAGTTTATGATGAACCTTAATAT	4979	XX	Human; pancreas; cancer; gene therapy; vaccine; immunostimulant;		
QY	1981	ATAATGTTCCAGGATTTAGTTCAATATTTGTTACTGTTATCTATCTGCTGATATAGG	2040	XX	cytostatic; tumour; gene; ss.		
Db	4980	ATAATGTTCCAGGATTTAGTTCAATATTTGTTACTGTTATCTATCTGCTGATATAGG	5039	KW	Homo sapiens.		
QY	2041	AAATCTTTTAAATCAAAAGCTGAAACGAATCAGCAATTTAGTCTTTCAGGACACACCCAA	2100	OS	WO200260317-A2.		
Db	5040	AAATCTTTTAAATCAAAAGCTGAAACGAATCAGCAATTTAGTCTTTCAGGACACACCCAA	5099	XX	08-AUG-2002.		
QY	2101	TAAATCAGTCAATGTAATGATGACAAAGTTTGTGTTTGTGTTTGTGTTTGTGTTGG	2160	XX	30-JAN-2002; 2002WO-US002781.		
Db	5100	TAAATCAGTCAATGTAATGATGACAAAGTTTGTGTTTGTGTTTGTGTTTGTGTTGG	5159	PR	30-JAN-2001; 2001US-0265305P.		
QY	2161	TTTTTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTGG	2220	PR	31-JAN-2001; 2001US-0265682P.		
Db	5160	TTTTTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTGG	5219	PR	09-FEB-2001; 2001US-0267568P.		
QY	2221	TAAAGGTTTGTAGTAAAGAAAGTCTCTGCTGATGATGATAGGGGCAAACTCTTTTC	2280	PR	21-MAR-2001; 2001US-0278651P.		
Db	5220	TAAAGGTTTGTAGTAAAGAAAGTCTCTGCTGATGATGATAGGGGCAAACTCTTTTC	5279	PR	28-APR-2001; 2001US-0287112P.		
QY	2281	CCCTTTCTGTTAATAGTCATCATTTCTATGCCAAACAGGAACGATCCATCTTGTAGT	2340	PR	16-MAY-2001; 2001US-0291631P.		
Db	5280	CCCTTTCTGTTAATAGTCATCATTTCTATGCCAAACAGGAACGATCCATCTTGTAGT	5339	PR	12-JUL-2001; 2001US-0305484P.		
QY	2341	CTTAATGTCACATTTGATTAATAATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTCG	2400	PR	20-AUG-2001; 2001US-0313999P.		
Db	5340	CTTAATGTCACATTTGATTAATAATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTCG	5399	PR	27-NOV-2001; 2001US-0333626P.		
QY	2401	TTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG	2460	XX	(CORI-) CORIXA CORP.		
Db	5400	TTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG	5459	XX	Benson DR, Kalos MD, Lodes MJ, Persing DH, Hepler WT, Jiang Y;		
QY	2461	GAAGCGTTGGGATCTTCAATTAATGATGAGGATCTGATGATGATGATGATGATGATGATG	2520	PI	WPI; 2002-627435/67.		
Db	5460	GAAGCGTTGGGATCTTCAATTAATGATGAGGATCTGATGATGATGATGATGATGATGATG	5519	XX	P-PSDB; ABP68622.		
QY	2521	CTGTGTTGGGCTCAACCGTCAATGTTGGGATGACGATGATGATGATGATGATGATGATG	2580	DR	New isolated polynucleotide and pancreatic tumor polypeptides, useful for		
Db	5520	CTGTGTTGGGCTCAACCGTCAATGTTGGGATGACGATGATGATGATGATGATGATGATG	5579	PT	diagnosing, preventing and/or treating cancer, particularly pancreatic		
				PS	cancer.		
				XX	Claim 1; SEQ ID NO 149; 300pp + Sequence Listing; English.		
				XX	The invention relates to an isolated polynucleotide (I) comprising: (a)		
				CC	any of a group of over 4000 nucleotide sequences (ABV94628-ABV99145); (b)		
				CC	complements of (a); (c) sequences consisting of at least 20 contiguous		
				CC	residues of (a); (d) sequences that hybridize to (a), under moderately		
				CC	stringent conditions; (e) sequences having at least 75% or 90% identity		
				CC	to (a); or (f) degenerate variants of (a). Polypeptides (ABP68596-		
				CC	ABP68637) encoded by (I) and oligonucleotide can be used to detect cancer		
				CC	in a patient and compositions comprising polypeptides, polynucleotides,		
				CC	antibodies, fusion proteins, T cell populations and antigen presenting		
				CC	cells expressing the polypeptide are useful in treating pancreatic cancer		
				CC	and stimulating an immune response. The polynucleotides can be used as		
				CC	probes or primers for nucleic acid hybridisation, in the design and		

RESULT 7	
ABQ77406	
ID	ABQ77406 standard; DNA; 5784 BP.
XX	
AC	ABQ77406;
XX	
DT	10-MAY-2003 (first entry)
XX	
DE	Human THBS2 DNA.
XX	
KW	Human; THBS2: vascular disease; cardiatic; antiarteriosclerotic; stroke;
KW	cerebroprotective; gene therapy; coronary artery disease; ischaemia;
KW	myocardial infarction; peripheral vascular disease; pulmonary embolism;
KW	venous thromboembolism; forensic; paternity testing; GI307505; gene; SNP;
KW	single nucleotide polymorphism; ds.
XX	
OS	Homo sapiens.
XX	

This invention describes a novel method for identifying a subject as a candidate for a particular course of therapy to treat a vascular disease or disorder. The method comprises determining the identity of the nucleotide present at specific positions, or their complements, and identifying the subject as a candidate for a particular clinical course of therapy based on the identity of the nucleotide present in that specific position. The method can be used for identifying a subject who is a candidate for further diagnostic evaluation of a vascular disease or disorder and selecting a clinical course of therapy. The products of the invention have cardiant, antiarteriosclerotic and cerebroprotective activity and can be used for gene therapy. The methods disclosed are useful for treating a vascular disease, e.g. atherosclerosis, coronary artery disease, myocardial infarction, ischaemia, stroke, peripheral vascular diseases, venous thromboembolism and pulmonary embolism. The DNA sequences are useful as fingerprint for detecting different individuals within the same species applicable in forensic studies and paternity testing. This sequence encodes the human THBS2 gene represented in [FIG. 1](#). Used to illustrate the method of the invention

QY	1	GAGGACTTGGACGGTGATGACGGGGTGATATTGTAAAGATGATTTTGGACAATGACCAAC	60
Db	3000	GAGGACTTGGACGGTGATGACGGGGTGATATTGTAAAGATGATTTTGGACAATGACCAAC	3059
QY	61	ATCCACAGATATTGATGATGTGTGTCCTGAAACAATGCCATCAGTGAGACAGACTTCAGG	120
Db	3060	ATCCACAGATATTGATGATGTGTGTCCTGAAACAATGCCATCAGTGAGACAGACTTCAGG	3119
QY	121	AACCTTCAGATGGTCCCTTGGATCCCAAAGGGACACGCCAAATTGATCCCACTGGGTC	180
Db	3120	AACCTTCAGATGGTCCCTTGGATCCCAAAGGGACACGCCAAATTGATCCCACTGGGTC	3179
QY	181	ATTTCGCATCAAGGCAAGGAGCTGGTTCAGACAGCCAACTCGGACCCCGCATCGCTGTA	240
Db	3180	ATTTCGCATCAAGGCAAGGAGCTGGTTCAGACAGCCAACTCGGACCCCGCATCGCTGTA	3239
QY	241	GGTTTTCAGGAGTTTGGGTCTGTGGACTTCAGTGCGCACATTCTACGTAAACACTGACCGG	300

Db 3240 GGTGTTGACGAGTTTGGGCTGTGGACTTCACTGAGGACATTTCTACGTAACACCTGACCGG 3299
QY 301 GACGACGACTATGCTGGCTTGGCTTTGGTTACCACTCAAGCAGCGCTTCTATGTTGGT 360
Db 3300 GACGACGACTATGCTGGCTTGGCTTTGGTTACCACTCAAGCAGCGCTTCTATGTTGGT 3359
QY 361 ATGTGAAGCAGGTGACGAGACCTACTGGGAGACCCAGCCACCGGGCCTATGCTATC 420
Db 3360 ATGTGAAGCAGGTGACGAGACCTACTGGGAGACCCAGCCACCGGGCCTATGCTATC 3419
QY 421 TCCGGGCTGCTCCCTCAAGTGTGAATCCACCA CGGGGACGGGAGCACTTGAGGAAC 480
Db 3420 TCCGGGCTGCTCCCTCAAGTGTGAATCCACCA CGGGGACGGGAGCACTTGAGGAAC 3479
QY 481 GCGCTGTGGCA CACGGGGAAACGCGGGGAGGTGCGAACCTTATGGCACGCCCCAGG 540
Db 3480 GCGCTGTGGCA CACGGGGAAACGCGGGGAGGTGCGAACCTTATGGCACGCCCCAGG 3539
QY 541 AACATTGGCTGAAGGACTACACGGCTTATAGGTGACACTGACTCACAGGCCCAAGACC 600
Db 3540 AACATTGGCTGAAGGACTACACGGCTTATAGGTGACACTGACTCACAGGCCCAAGACC 3599
QY 601 GGCTACATCAGAGTCTTAGTGCATGAAGAAACAGGTCATGGCAGACTCAGGACCTATC 660
Db 3600 GGCTACATCAGAGTCTTAGTGCATGAAGAAACAGGTCATGGCAGACTCAGGACCTATC 3659
QY 661 TATGACCAAACTACGCTGGCGGGCGGCTGCTATTTGCTCTCTCAAGAAATGGTC 720
Db 3660 TATGACCAAACTACGCTGGCGGGCGGCTGCTATTTGCTCTCTCAAGAAATGGTC 3719
QY 721 TATTTCTCAGACCTCAAGTACGAATGACAGATAATTTAAACAAGATTGCTGCAATTCGG 780
Db 3720 TATTTCTCAGACCTCAAGTACGAATGACAGATAATTTAAACAAGATTGCTGCAATTCGG 3779
QY 781 GCATGCGCTGTCATGTCATGGTCCCTAGACACCTCAGTTCATTTGGTGGTCTTCGGGCT 840
Db 3780 GCATGCGCTGTCATGTCATGGTCCCTAGACACCTCAGTTCATTTGGTGGTCTTCGGGCT 3839
QY 841 TCTCTCTCAGCAGCCTCTGCTCCCTTGACCTTAACTCTGATGGTCTTCACTCCTG 900
Db 3840 TCTCTCTCAGCAGCCTCTGCTCCCTTGACCTTAACTCTGATGGTCTTCACTCCTG 3899
QY 901 CCAGCAACCCCAACCCCAAGTCCCTCAGAGATAAATATCAATGGAATCAGAGATGAA 960
Db 3900 CCAGCAACCCCAACCCCAAGTCCCTCAGAGATAAATATCAATGGAATCAGAGATGAA 3959
QY 961 CATCTAACCCACTAGAGAAACCAAGTTGGTGATATGAGACTTTATGGAGTGAAAA 1020
Db 3960 CATCTAACCCACTAGAGAAACCAAGTTGGTGATATGAGACTTTATGGAGTGAAAA 4019
QY 1021 TTGGGCATGCCATTACATTTGCTTTTCTGTTTGTGTTTAAAGAAATGAGCTTTACATATA 1080
Db 4020 TTGGGCATGCCATTACATTTGCTTTTCTGTTTGTGTTTAAAGAAATGAGCTTTACATATA 4079
QY 1081 AAATGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1140
Db 4080 AAATGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4139
QY 1141 CCATATGATATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1200
Db 4140 CCATATGATATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4199
QY 1201 CACTATTTCTTGAAATTAGAGTTGCTCAATATGACACACCAATCCCGCTTAATAATTAATA 1260
Db 4200 CACTATTTCTTGAAATTAGAGTTGCTCAATATGACACACCAATCCCGCTTAATAATTAATA 4259
QY 1261 AACAGGGTCAATTTCAATTTGAAGTAATGTTTGAAGGAGAGATTTAGAGACCAACAG 1320
Db 4260 AACAGGGTCAATTTCAATTTGAAGTAATGTTTGAAGGAGAGATTTAGAGACCAACAG 4319
QY 1321 GCATAGCAATGACATTAAGCTACCGATTAACTAATCGGAACATGTAATAACAGTTTACAAA 1380

Db 4320 GCATAGCAATGACATTAAGCTACCGATTAACTAATCGGAACATGTAATAACAGTTACAAA 4379
QY 1381 ATAAACGAACTCTCTCTTGTGCTCAATGAAGCCCTCATGTGCAGTAGAGATGAGATT 1440
Db 4380 ATAAACGAACTCTCTCTTGTGCTCAATGAAGCCCTCATGTGCAGTAGAGATGAGATT 4439
QY 1441 TCATCAAAAGAAACAAACATCCTTGGCAATGGGTGTGACGCGGTTCCAGATGTGGAATGGC 1500
Db 4440 TCATCAAAAGAAACAAACATCCTTGGCAATGGGTGTGACGCGGTTCCAGATGTGGAATGGC 4499
QY 1501 AAAACCTCATTTAAAGTAAAGGTTAGCAGACGAAGGTGGGTGCTTTAGCTGCTGCTTCT 1560
Db 4500 AAAACCTCATTTAAAGTAAAGGTTAGCAGACGAAGGTGGGTGCTTTAGCTGCTGCTTCT 4559
QY 1561 GCGGTTGGGCTCGGGGAGGCTCTGCTGAGCTTCCCTCCAGCTTTGCTGCTGCTGAG 1620
Db 4560 GCGGTTGGGCTCGGGGAGGCTCTGCTGAGCTTCCCTCCAGCTTTGCTGCTGCTGAG 4619
QY 1621 AGGAACCCAGACGACGACGCGGCGGAAAGGCGCATCTAACGGGTATCTAGGCTTTGG 1680
Db 4620 AGGAACCCAGACGACGACGCGGCGGAAAGGCGCATCTAACGGGTATCTAGGCTTTGG 4679
QY 1681 TAACTGCGGACAAAGTTGCTTTTACCTGATTTGATGATACATTTCAATTAAGGTTCCAGTTA 1740
Db 4680 TAACTGCGGACAAAGTTGCTTTTACCTGATTTGATGATACATTTCAATTAAGGTTCCAGTTA 4739
QY 1741 TAAATATTTTGTATATTTATTAAGTGACATATAGAATGCAACTCCATTTACAGTAAC 1800
Db 4740 TAAATATTTTGTATATTTATTAAGTGACATATAGAATGCAACTCCATTTACAGTAAC 4799
QY 1801 TATTTTAAATATGCTAGTAAACATATAGTAGTATAATTTCTAGAAAACAAACATCTAATA 1860
Db 4800 TATTTTAAATATGCTAGTAAACATATAGTAGTATAATTTCTAGAAAACAAACATCTAATA 4859
QY 1861 AGTATATAATCTGTGAAATATAGAGCTTGATATAATTTAGTGTGTACGATGAAGCATG 1920
Db 4860 AGTATATAATCTGTGAAATATAGAGCTTGATATAATTTAGTGTGTACGATGAAGCATG 4919
QY 1921 CTAGAAGCTGTAAACAGAAATACATAGAGAAATATAGAGAGTTTATGATGGAACCTTTAATAT 1980
Db 4920 CTAGAAGCTGTAAACAGAAATACATAGAGAAATATAGAGAGTTTATGATGGAACCTTTAATAT 4979
QY 1981 ATAATGTTGCGCAGGATTTAGTTCAATATTTGTTTCTGTTTATCTATCTGCTGTATATGG 2040
Db 4980 ATAATGTTGCGCAGGATTTAGTTCAATATTTGTTTCTGTTTATCTATCTGCTGTATATGG 5039
QY 2041 AATTTCTTTTAAATCAAAACGCTGAAACGAATCAGCATTTAGTCTTGCCAGGACACCCAA 2100
Db 5040 AATTTCTTTTAAATCAAAACGCTGAAACGAATCAGCATTTAGTCTTGCCAGGACACCCAA 5099
QY 2101 TAATCAGTCACTGTATATATGACAAAGTTGTTTGTGTTTGTGTTTGTGTTGTTGGTGG 2160
Db 5100 TAATCAGTCACTGTATATATGACAAAGTTGTTTGTGTTTGTGTTTGTGTTTGTGTTGG 5159
QY 2161 TTTTGTGCTTTAAGTTGCAATGATCTTCTGCGAGAAATAGTCACTATCCACTCCACA 2220
Db 5160 TTTTGTGCTTTAAGTTGCAATGATCTTCTGCGAGAAATAGTCACTATCCACTCCACA 5219
QY 2221 TAAGGGGTTTAAAGGAAAGTCTGCTGTCTGATGATGATAGGGGCAAACTCTTTTC 2280
Db 5220 TAAGGGGTTTAAAGGAAAGTCTGCTGTCTGATGATGATAGGGGCAAACTCTTTTC 5279
QY 2281 CCCTTTCTGTTAATAGTCACTACATTTCTATGCAAAACAGAAACGATCCATACTTTAGT 2340
Db 5280 CCCTTTCTGTTAATAGTCACTACATTTCTATGCAAAACAGAAACGATCCATACTTTAGT 5339
QY 2341 CTTAAATGTAACATTCGATTTGATATAAATTAATTTGTTTCTTCTGAGGTTGATCG 2400
Db 5340 CTTAAATGTAACATTCGATTTGATATAAATTAATTTGTTTCTTCTGAGGTTGATCG 5399
QY 2401 TTGTTGTTTGTGCTGACATTTTACTTTTTTTCGGTGTGGAGCTGATTTCCCGAGCAAC 2460
Db 5400 TTGTTGTTTGTGCTGACATTTTACTTTTTTTCGGTGTGGAGCTGATTTCCCGAGCAAC 5459

QY 2461 GAAGCGTTGGGATACCTTCATTAATGATGAGCTGTCAACAGCGGTGACGGTTTCTGTTT 2520
Db |||||||
QY 5460 GAAGCGTTGGGATACCTTCATTAATGATGAGCTGTCAACAGCGGTGACGGTTTCTGTTT 5519
Db |||||||
QY 2521 CTGTGTTGTGGGTCAACCGTACCAATGCTGGGAATGACGATGATGGAATTTAGAA 2580
Db |||||||
QY 5520 CTGTGTTGTGGGTCAACCGTACCAATGCTGGGAATGACGATGATGGAATTTAGAA 5579
Db |||||||
QY 2581 TGACACATATTTTGTAAATTTATGTTTCTTAAACAAATTTATCTATAGGTTGA 2640
Db |||||||
QY 5580 TGTACCATATTTTGTAAATTTATGTTTCTTAAACAAATTTATCTATAGGTTGA 5639
Db |||||||
QY 2641 TGAACGTCATGTTTGTGCAAGACTGTAATATTTATTTATGTTTCAATGTTCA 2700
Db |||||||
QY 5640 TGAACGTCATGTTTGTGCAAGACTGTAATATTTATTTATGTTTCAATGTTCA 5699
Db |||||||
QY 2701 AATTTACCACTGAACCCCTGCACCTAGCTAGAGAACCTCATTTTAAAGATTAAACAGG 2760
Db |||||||
QY 5700 AATTTACCACTGAACCCCTGCACCTAGCTAGAGAACCTCATTTTAAAGATTAAACAGG 5759
Db |||||||
QY 2761 AATTAATTTGTAAGGTTTCT 2785
Db |||||||
QY 5760 AATTAATTTGTAAGGTTTCT 5784
Db |||||||

RESULT 8

AAD55115
ID AAD55115 standard; DNA; 5784 BP.
XX
AC AAD55115;
XX
DT 07-AUG-2003 (first entry)
XX
DE Human THB2 reference DNA (GI 307505).
XX
KW Thrombospondin 2; THB2; angiotensin converting enzyme; polymorphism;
KW ACE-1; beta-fibrinogen; FGB; peripheral vascular disease; ischaemia;
KW vascular disease; myocardial infarction; pulmonary embolism; stroke;
KW atherosclerosis; coronary artery disease; venous thromboembolism; human;
KW gene; ds.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 240..3758
FT /*tag= a
FT /product= "Human THB2 reference protein"
FT replace(3949,G)
FT /*tag= b
FT /standard_name= "Single nucleotide polymorphism (SNP)"
XX
XX WO2003020118-A2.
XX
XX 13-MAR-2003.
XX
XX 04-SEP-2002; 2002WO-US028113.
XX
XX 05-SEP-2001; 2001US-0317178P.
XX 16-OCT-2001; 2001US-0329958P.
XX 14-DEC-2001; 2001US-0001772A.
XX
XX (VITI-) VITIITY INC.
XX
XX Mccarthy J;
XX
XX WPI; 2003-300816/29.
XX P-PSDB; AAE36411.
XX
XX Identifying polymorphisms in thrombospondin 2, angiotensin converting
XX enzyme and/or beta-fibrinogen genes in nucleic acid sample of subject, by
XX contacting the nucleic acid with a complementary probe or primer.
XX

Claim 2; Fig 1; 194pp; English.

PS The invention relates to a method for determining the identity of one or
XX more allelic variants of a polymorphic region of a thrombospondin 2
CC (THB2), angiotensin converting enzyme (ACE)-1 and/or beta-fibrinogen
CC (FGB) genes in a nucleic acid obtained from a subject. The method
CC involves contacting the nucleic acid with a complementary probe or
CC primer. The method is useful for diagnosing or aiding in the diagnosis of
CC coronary artery disease, atherosclerosis, ischaemia, stroke, peripheral
CC vascular disease, venous thromboembolism and pulmonary embolism. The
CC present sequence is human THB2 reference DNA
XX

SQ Sequence 5784 BP; 1447 A; 1460 C; 1518 G; 1359 T; 0 U; 0 Other;

Query Match 100.0%; Score 2785; DB 9; Length 5784;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2785; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GAGGACTTGGACGGTGTGACGGGGTATTTTAAAGATGATTTTGAATGACATGACAA 60
Db |||||||
QY 61 ATCCAGATATTGATGATGTGTCTGAAACAAATGCCATCAGTGAGACAGACTTCAGG 120
Db |||||||
QY 3060 ATCCAGATATTGATGATGTGTCTGAAACAAATGCCATCAGTGAGACAGACTTCAGG 3119
Db |||||||
QY 121 AACTTCCAGATGTCTCCCTTTGGATCCCAAGGGACCAACCCAAATGTATCCCACTGGGTC 180
Db |||||||
QY 3120 AACTTCCAGATGTCTCCCTTTGGATCCCAAGGGACCAACCCAAATGTATCCCACTGGGTC 3179
Db |||||||
QY 181 ATTGCCATCAAGCAAGGAGCTGTTTCAGACAGCAACTCGGACCCGGCATCGCTGTA 240
Db |||||||
QY 3180 ATTGCCATCAAGCAAGGAGCTGTTTCAGACAGCAACTCGGACCCGGCATCGCTGTA 3239
Db |||||||
QY 241 GGTTTTCAGGAGTTTGGTCTGTGCACTTCACTGGGCACTTACGTAAACACATGACCGG 300
Db |||||||
QY 3240 GGTTTTCAGGAGTTTGGTCTGTGCACTTCACTGGGCACTTACGTAAACACATGACCGG 3299
Db |||||||
QY 301 GACGACGACTATGCTGGCTTCGTCTTTGGTTACAGTCAAGCAGCGCTTCTATGTGTG 360
Db |||||||
QY 3300 GACGACGACTATGCTGGCTTCGTCTTTGGTTACAGTCAAGCAGCGCTTCTATGTGTG 3359
Db |||||||
QY 361 ATGTGGAAGAGGTGACGACACCTACTGCGAGGACAGCCACCGCGGCTTATGGCTAC 420
Db |||||||
QY 3360 ATGTGGAAGAGGTGACGACACCTACTGCGAGGACAGCCACCGCGGCTTATGGCTAC 3419
Db |||||||
QY 421 TCCGCGGTGTCCTCAAGGTGCTGAACTCCACACCGGGGACGGCGAGCACCTGAGGAAC 480
Db |||||||
QY 3420 TCCGCGGTGTCCTCAAGGTGCTGAACTCCACACCGGGGACGGCGAGCACCTGAGGAAC 3479
Db |||||||
QY 481 GCGCTGTGGCAACCGGGGAAACACCGCGGGGAGGTGCGAACTTATGGCAGACCCCGAG 540
Db |||||||
QY 3480 GCGCTGTGGCAACCGGGGAAACACCGCGGGGAGGTGCGAACTTATGGCAGACCCCGAG 3539
Db |||||||
QY 541 AACATGCTGGAAGACTTACACGCGCTATAGTGGCACTTACGTCAAGGCCCCAAGACC 600
Db |||||||
QY 3540 AACATGCTGGAAGACTTACACGCGCTATAGTGGCACTTACGTCAAGGCCCCAAGACC 3599
Db |||||||
QY 601 GGCTACATCAGAGTCTTATGTCATGAAGGAAACAGGTCATGGCAGACTCAGGACCTATC 660
Db |||||||
QY 3600 GGCTACATCAGAGTCTTATGTCATGAAGGAAACAGGTCATGGCAGACTCAGGACCTATC 3659
Db |||||||
QY 661 TATGACCAAACTTACGCTGGCGGGCTGTTGTTCTTCTCTCAAGAAATGGTC 720
Db |||||||
QY 3660 TATGACCAAACTTACGCTGGCGGGCTGTTGTTCTTCTCTCAAGAAATGGTC 3719
Db |||||||
QY 721 TATTTCTCAGACTCAAGTACGAAATGACAGATATTTAAACAAGATTTCTGCTCATTTCCG 780
Db |||||||
QY 3720 TATTTCTCAGACTCAAGTACGAAATGACAGATATTTAAACAAGATTTCTGCTCATTTCCG 3779
Db |||||||
QY 781 GCAATGCCCTGTGATGCCATCGTCCCTAGACACCTCAGTTCATTTGTTGCTTGGGCT 840
Db |||||||

D _b	3780	GCAATGCCCTGFCATGCCATGGTCCCTAGACACTCAGTTTCATGTGGTGCTTCTCGCGCT	3831
Q _y	841	TCTCTCTCTAGCAGCACCTCTGTCCCTTGACCCTTAACCTCTGATGGTTCTTCACCTCTCG	900
D _b	3840	TCTCTCTCTAGCAGCACCTCTGTCCCTTGACCCTTAACCTCTGATGGTTCTTCACCTCTCG	3899
Q _y	901	CCAGCAACCCCAAACCCCAAGTGCCTTCAGAGGATAAATATCAATGGAACTCAGAGATGAA	960
D _b	3900	CCAGCAACCCCAAACCCCAAGTGCCTTCAGAGGATAAATATCAATGGAACTCAGAGATGAA	3959
Q _y	961	CATCTAACCCCACTAGAGAAACCCAGTTTGGTGATATATGAGACTTTATGTGGAGTGAAAA	1020
D _b	3960	CATCTAACCCCACTAGAGAAACCCAGTTTGGTGATATATGAGACTTTATGTGGAGTGAAAA	4019
Q _y	1021	TTGGGCATGCCATTACATTGCTTTTTCTGTGTGTTTAAAAGAATCAGCGTTTACATATA	1080
D _b	4020	TTGGGCATGCCATTACATTGCTTTTTCTGTGTGTTTAAAAGAATCAGCGTTTACATATA	4079
Q _y	1081	AAATGTAATTACTTATTGTATTTATGTATATGAGTTGAAGGGAATACTGTGCATAAG	1140
D _b	4080	AAATGTAATTACTTATTGTATTTATGTATATGAGTTGAAGGGAATACTGTGCATAAG	4139
Q _y	1141	CCATTATGATAAATTAAGCATGAAAAATATTCGTAACCTATTTTGGTGCTTAAAGTTGT	1200
D _b	4140	CCATTATGATAAATTAAGCATGAAAAATATTCGTAACCTATTTTGGTGCTTAAAGTTGT	4199
Q _y	1201	CACTATTTCTGAAATTAGAGTTGCTCTACAATGACACACAACAAATCCGCTAATAAATTATA	1260
D _b	4200	CACTATTTCTGAAATTAGAGTTGCTCTACAATGACACACAACAAATCCGCTAATAAATTATA	4259
Q _y	1261	AACHAGGTCATTTCAAATTTGAAGTAATGTTTTTAGTAGGAGAGATTGAAGACAAACAG	1320
D _b	4260	AACHAGGTCATTTCAAATTTGAAGTAATGTTTTTAGTAGGAGAGATTGAAGACAAACAG	4319
Q _y	1321	GCATAGCAAAATGACATAAGCTACCGATTAACCTAAATCGGAACTGTGTAACAGTTACAAA	1380
D _b	4320	GCATAGCAAAATGACATAAGCTACCGATTAACCTAAATCGGAACTGTGTAACAGTTACAAA	4379
Q _y	1381	ATAAACGAACTCTCTCTGTCTTACAATCAAAAGCCCTCATGTGCGAGTAGAGTGCAGTT	1440
D _b	4380	ATAAACGAACTCTCTCTGTCTTACAATCAAAAGCCCTCATGTGCGAGTAGAGTGCAGTT	4439
Q _y	1441	TCATCAAGAACAAACATCCTTGCAAAATGGGTGTGACGGGTTCCAGATGTGGATTGGC	1500
D _b	4440	TCATCAAGAACAAACATCCTTGCAAAATGGGTGTGACGGGTTCCAGATGTGGATTGGC	4499
Q _y	1501	AAAACCTCATTTAAGTAAAGGTTAGCAGACAAAGTCGGGTGCTTTAGCTGCTGTGT	1560
D _b	4500	AAAACCTCATTTAAGTAAAGGTTAGCAGACAAAGTCGGGTGCTTTAGCTGCTGTGT	4559
Q _y	1561	GCGGTTGTGGCGTGGGGGCTCCTGCCTGAGCTTCTTTCCACAGCTTTGCTGCTGAG	1620
D _b	4560	GCGGTTGTGGCGTGGGGGCTCCTGCCTGAGCTTCTTTCCACAGCTTTGCTGCTGAG	4619
Q _y	1621	AGGAACAGAGCAGACGCA CAGCGCGGAAAGCGCATCTAACGGTATCTAGGCTTTGG	1680
D _b	4620	AGGAACAGAGCAGACGCA CAGCGCGGAAAGCGCATCTAACGGTATCTAGGCTTTGG	4679
Q _y	1681	TAACTCGGCAAGTTGCTTTTACCTGATTGTGATGATCAATTTCAATAGGTTCCAGTTA	1740
D _b	4680	TAACTCGGCAAGTTGCTTTTACCTGATTGTGATGATCAATTTCAATAGGTTCCAGTTA	4739
Q _y	1741	TAAATATTTTGTAAATTTTAAAGTGACTATAGAAATGCAACTCCATTTACCACTACT	1800
D _b	4740	TAAATATTTTGTAAATTTTAAAGTGACTATAGAAATGCAACTCCATTTACCACTACT	4799
Q _y	1801	TATTTTAAATATGCCTAGTAAACATATGTAGTATATAATTTCTAGAAAACAAACATCTAATA	1860
D _b	4800	TATTTTAAATATGCCTAGTAAACATATGTAGTATATAATTTCTAGAAAACAAACATCTAATA	4859
Q _y	1861	AGTATATATCTCTGAAATAATAGCGCTTGATATATTTAGTGTGTGACGATGAAGCATG	1920
D _b	4860	AGTATATATCTCTGAAATAATAGCGCTTGATATATTTAGTGTGTGACGATGAAGCATG	4919

QY	1921	CTAGAAGCTGTAAACAGAAATACATAGAGAAATAATGAGGAGTTTATGATGGAACCTTAATAT	1980
Db	4920	CTAGAAGCTGTAAACAGAAATACATAGAGAAATAATGAGGAGTTTATGATGGAACCTTAATAT	4979
QY	1981	ATAAATGTTGCCAGCGATTTTAGTTCAATATTTGTTACTGTATCTACTCTGTATATGG	2040
Db	4980	ATAAATGTTGCCAGCGATTTTAGTTCAATATTTGTTACTGTATCTACTCTGTATATGG	5039
QY	2041	AATTCCTTTTAATTCAAACGCTGAAAAAGAAATCAGCATTTAGTCTTCCAGGCGACACCCAA	2100
Db	5040	AATTCCTTTTAATTCAAACGCTGAAAAAGAAATCAGCATTTAGTCTTCCAGGCGACACCCAA	5099
QY	2101	TAATCAGTCATGTGTAAATATGACAAGTTTGTCTTTTCTTTTGTGTTTTTGTGTTGTTGGTGG	2160
Db	5100	TAATCAGTCATGTGTAAATATGACAAGTTTGTCTTTTCTTTTGTGTTTTTGTGTTGTTGGTGG	5159
QY	2161	TTTTTTTGCCTTAAGTTTGCATGATCTTTCTGCAGGAATAAGTCACCTCATCCCACTCCACA	2220
Db	5160	TTTTTTTGCCTTAAGTTTGCATGATCTTTCTGCAGGAATAAGTCACCTCATCCCACTCCACA	5219
QY	2221	TAAAGGGTGTAGTAAGAGAAAGTCGTCTGTCTCATGATGGAATAGGCGGGAATCTTTTTTC	2280
Db	5220	TAAAGGGTGTAGTAAGAGAAAGTCGTCTGTCTCATGATGGAATAGGCGGGAATCTTTTTTC	5279
QY	2281	CCCTTTCTGTAAATAGTCATCACATTTCTATGCCAACAGGAACGATCCATACTTTAGT	2340
Db	5280	CCCTTTCTGTAAATAGTCATCACATTTCTATGCCAACAGGAACGATCCATACTTTAGT	5339
QY	2341	CTTAATGTACACATGTGCATTTTGATAAAATTAATTTTGTGTTTCCCTTTGAGGTTGATCG	2400
Db	5340	CTTAATGTACACATGTGCATTTTGATAAAATTAATTTTGTGTTTCCCTTTGAGGTTGATCG	5399
QY	2401	TTGTGTTGTTTTGCTGCACATTTTACTCTTTTTCGGTGTGAGCTGTATTTCCCGAGACAAC	2460
Db	5400	TTGTGTTGTTTTGCTGCACATTTTACTCTTTTTCGGTGTGAGCTGTATTTCCCGAGACAAC	5459
QY	2461	GAAGCGTGTGGGATCTTCATTAATGTAGCGACTGTCAACAGCGTGCAGGTTTCTCGTTT	2520
Db	5460	GAAGCGTGTGGGATCTTCATTAATGTAGCGACTGTCAACAGCGTGCAGGTTTCTCGTTT	5519
QY	2521	CTGTGTTGTGGGTCACCGGTACAAATGGTGTGGGAATGACGATGATGAATATTTAGAA	2580
Db	5520	CTGTGTTGTGGGTCACCGGTACAAATGGTGTGGGAATGACGATGATGAATATTTAGAA	5579
QY	2581	TGTATCCATATTTTTGTAAATATTTATGTTTTTCTAAACAAATTTATCGTATAGTTGA	2640
Db	5580	TGTATCCATATTTTTGTAAATATTTATGTTTTTCTAAACAAATTTATCGTATAGTTGA	5639
QY	2641	TGAAACGTCTATGTTTTGTCAAAGACTGTAAATATTTTATTTATGTTTACATCGTCA	2700
Db	5640	TGAAACGTCTATGTTTTGTCAAAGACTGTAAATATTTTATTTATGTTTACATCGTCA	5699
QY	2701	AATTTACCACTGAAACCCCTGCATTTAGCTAGAACCTCATTTTTTAAAGATTAAACAACAGG	2760
Db	5700	AATTTACCACTGAAACCCCTGCATTTAGCTAGAACCTCATTTTTTAAAGATTAAACAACAGG	5759
QY	2761	AAATAAATTTGTAATAAGGTTTTCT	2785
Db	5760	AAATAAATTTGTAATAAGGTTTTCT	5784

RESULT 9
ADD31094
ID ADD31094 standard; cDNA; 5784 BP.

DT 15-JAN-2004 (first entry)

DE Human cDNA encoding thrombospondin 2, THBS2.

Human; thrombospondin 2; THBS2; ss; gene; single nucleotide polymorphism;
KW

SNP: antiarteriosclerotic; cardiant; vasotropic; cerebroprotective;
cardiovascular-Gen; Thrombolytic; gene therapy; atherosclerosis;
coronary heart disease; myocardial infarction; stroke;
peripheral vascular diseases; venous thromboembolism; pulmonary embolism;
chromosome 6q27.

Homo sapiens.

Key Location/Qualifiers
CDS 240..3758
FT /*tag= b
FT /product= "Thrombospondin 2"
FT sig_peptide 240..293
FT /*tag= a
FT mat_peptide 294..3754
FT /*tag= c
FT /note= "Mature thrombospondin 2"
FT variation replace(3949,G)
FT /*tag= d
FT /standard_name= "Single nucleotide polymorphism"
FT variation replace(5489,T)
FT /*tag= e
FT /standard_name= "Single nucleotide polymorphism"

US2003194703-A1.

16-OCT-2003.

13-NOV-2001; 2001US-00007781.

13-NOV-2000; 2000US-0248130P.

22-JUN-2001; 2001US-0300158P.

(BOLK/) BOLK S.

(DALEY/) DALEY G. Q.

(MCCA/) MCCARTHY J. J.

Bolk S, Daley GO, McCarthy JJ;

WPI; 2003-844447/78.

P-PSDB; ADD31095.

Predicting the likelihood of a vascular disease (e.g. stroke) in an individual comprises obtaining a nucleic acid sample from the individual and detecting a single nucleotide polymorphisms in the thrombospondin-2 gene.

Claim 2: SEQ ID NO 1; 17pp; English.

The invention relates to predicting the likelihood of a vascular disease, comprising obtaining a nucleic acid sample from the individual, and determining the genotype of the individual at nucleotide position 3949 of the thrombospondin-2 (THBS2) gene, where an individual who is homozygous for the variant allele has a decreased likelihood of a vascular disease as compared with an individual who is heterozygous or homozygous for the reference allele. Also included are diagnosing or aiding in the diagnosis of a vascular disease in an individual (comprising obtaining a nucleic acid sample from the individual, and determining the nucleotide present at nucleotide position 3949 of the thrombospondin-2 gene, where the presence of a T at nucleotide 3949 is indicative of an increased likelihood of a vascular disease in the individual, as compared with an individual having G at position 3949), a nucleic acid molecule comprising all or a portion of a sequence appearing as ADD31094 (where the nucleic acid molecule is at least 10 nucleotides in length and where the nucleic acid sequence comprises a polymorphic site at nucleotide position 3949 of ADD31094) and an allele-specific oligonucleotide that hybridizes to the nucleic acid molecule cited above. The composition and methods are useful in diagnosing, treating or predicting the clinical course and treatment response of vascular diseases (e.g. atherosclerosis, coronary heart disease, myocardial infarction, stroke, peripheral vascular diseases, venous thromboembolism or pulmonary embolism). The nucleic acid may also be used in developing new treatments for vascular disease and in developing cell culture-based and animal models for research and

CC treatment of the disease. The gene for THBS2 is located on chromosome
CC 6q27. The present sequence is the cDNA encoding thrombospondin 2.
XX
SQ Sequence 5784 bp; 1447 A; 1460 C; 1518 G; 1359 T; 0 U; 0 Other;
Query Match 100.0%; Score 2785; DB 10; Length 5784;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2785; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GAGGACTTGGACGCTGATGACCGGGGTGATATTTGTAAGATGATTTTGAACAAC 60
DB 3000 GAGGACTTGGACGCTGATGACCGGGGTGATATTTGTAAGATGATTTTGAACAAC 3059
QY 61 ATCCAGATATTGATGATGATGCTCTGTAAGAAACAATGCCATCAGTGAGACAGCTT 120
DB 3060 ATCCAGATATTGATGATGATGCTCTGTAAGAAACAATGCCATCAGTGAGACAGCTT 3119
QY 121 AACTTCCAGATGTCCTTGGATCCCAAGGACCAACCAAAATTTGATCCCACTGGTGC 180
DB 3120 AACTTCCAGATGTCCTTGGATCCCAAGGACCAACCAAAATTTGATCCCACTGGTGC 3179
QY 181 ATTGCCCATCAAGCAAGGAGCTGTTTCAGACAGCAACCTCGGACCCCGCATCGCTGTA 240
DB 3180 ATTGCCCATCAAGCAAGGAGCTGTTTCAGACAGCAACCTCGGACCCCGCATCGCTGTA 3239
QY 241 GGTTCGACGATTTGGTCTGTGGACTTCAGTGGCACTTCTACCTAATGACATGACCGG 300
DB 3240 GGTTCGACGATTTGGTCTGTGGACTTCAGTGGCACTTCTACCTAATGACATGACCGG 3299
QY 301 GACGACGACTATGCTGCTTCTTGTGTACAGTCAAGCAGCCGCTTCTATGTGGTG 360
DB 3300 GACGACGACTATGCTGCTTCTTGTGTACAGTCAAGCAGCCGCTTCTATGTGGTG 3359
QY 361 ATGTGGAAGCAGGTGACGACGACCTTCTGAGGAGCAACGACCCGCTTATGGCTAC 420
DB 3360 ATGTGGAAGCAGGTGACGACGACCTTCTGAGGAGCAACGACCCGCTTATGGCTAC 3419
QY 421 TCCGGCGTGTCCCTCAAGGTGGTGAATCTCCACCAAGGAGCGGCGGAGCCTGAGGAA 480
DB 3420 TCCGGCGTGTCCCTCAAGGTGGTGAATCTCCACCAAGGAGCGGCGGAGCCTGAGGAA 3479
QY 481 GCGTGTGGCACGCGGGAACACGCGGGCAGGTGCGAACCTTATGCAACGACCCGAGG 540
DB 3480 GCGTGTGGCACGCGGGAACACGCGGGCAGGTGCGAACCTTATGCAACGACCCGAGG 3539
QY 541 AACATTGCTGGAAGGACTACACGCGCTATAGTGGCAGCTGACATCAAGGCCCAAGACC 600
DB 3540 AACATTGCTGGAAGGACTACACGCGCTATAGTGGCAGCTGACATCAAGGCCCAAGACC 3599
QY 601 GGCTACATCAGAGTCTTAGTGTGATGAAGAAACAGGTCTATGGCAGACTCAGGACTATC 660
DB 3600 GGCTACATCAGAGTCTTAGTGTGATGAAGAAACAGGTCTATGGCAGACTCAGGACTATC 3659
QY 661 TATGACCAAACTACGCTGGCGGGCGGTGGTCTTATTTGCTTCTCTCAAGAAATGGTGC 720
DB 3660 TATGACCAAACTACGCTGGCGGGCGGTGGTCTTATTTGCTTCTCTCAAGAAATGGTGC 3719
QY 721 TATTTCTCAGACCTCAAGTACGAATGCAAGATATTTAAACAAGATTTGCTGCATTCGG 780
DB 3720 TATTTCTCAGACCTCAAGTACGAATGCAAGATATTTAAACAAGATTTGCTGCATTCGG 3779
QY 781 GCAATGCCCTGTGATGTCATGGTCCCTAGACACCTCAGTTCATTTGCTGCTCGGCT 840
DB 3780 GCAATGCCCTGTGATGTCATGGTCCCTAGACACCTCAGTTCATTTGCTGCTCGGCT 3839
QY 841 TCTCTCTTAGCAGACCTCTCTGCTCCCTTGAACCTTAACTCTGATGGTCTTCACTCTCTG 900
DB 3840 TCTCTCTTAGCAGACCTCTCTGCTCCCTTGAACCTTAACTCTGATGGTCTTCACTCTCTG 3899
QY 901 CCAGCAACCCCAAAACCCAAAGTGCCTTCAGAGGATTAATCAATGAATCAGAGTGA 960
DB 3900 CCAGCAACCCCAAAACCCAAAGTGCCTTCAGAGGATTAATCAATGAATCAGAGTGA 3959

961 CATCTAACCCACTAGAGGAAACCCAGTTTGGTGATATATGAGACTTTATGTGGAGTCAAAA 1020
Db |||||
3960 CATCTAACCCACTAGAGGAAACCCAGTTTGGTGATATATGAGACTTTATGTGGAGTCAAAA 4019
Qy |||||
1021 TTGGCGATGCCAATTACATTCCTTTTCTTTGTTTAAAGAAAGATGAGCTTTACATATA 1080
Db |||||
4020 TTGGCGATGCCAATTACATTCCTTTTCTTTGTTTAAAGAAAGATGAGCTTTACATATA 4079
Qy |||||
1081 AAATGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1140
Db |||||
4080 AAATGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4139
Qy |||||
1141 CAATTAATGATAAATTAAGCATGAAATAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1200
Db |||||
4140 CCAATTAATGATAAATTAAGCATGAAATAATTAATTAATTAATTAATTAATTAATTAATTAAT 4199
Qy |||||
1201 CACTATTCTTGAATTAAGTTGCTCTCAATGACACACAATTAATTAATTAATTAATTAATTAAT 1260
Db |||||
4200 CACTATTCTTGAATTAAGTTGCTCTCAATGACACACAATTAATTAATTAATTAATTAATTAAT 4259
Qy |||||
1261 AACAGGGTCAATTCAAATTTCAAGTAAATTTAGTAAAGGGAATTAATTAATTAATTAATTAAT 1320
Db |||||
4260 AACAGGGTCAATTCAAATTTCAAGTAAATTTAGTAAAGGGAATTAATTAATTAATTAATTAAT 4319
Qy |||||
1321 GCATAGCAAAATGACATAAGCTACCGATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1380
Db |||||
4320 GCATAGCAAAATGACATAAGCTACCGATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4379
Qy |||||
1381 ATAAAGCAATCTCTCTGCTCTCAATTAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAA 1440
Db |||||
4380 ATAAAGCAATCTCTCTGCTCTCAATTAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAA 4439
Qy |||||
1441 TCATCAAAAGAAACAAATCTCTGCTCTCAATTAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAA 1500
Db |||||
4440 TCATCAAAAGAAACAAATCTCTGCTCTCAATTAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAA 4499
Qy |||||
1501 AAAACCTCAATTAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAA 1560
Db |||||
4500 AAAACCTCAATTAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAA 4559
Qy |||||
1561 GCCGTTGGGCTGGGAGGCTCTGCTGAGCTTCCCTCCAGCTTTCTCCAGCTTTCTGCTGAG 1620
Db |||||
4560 GCCGTTGGGCTGGGAGGCTCTGCTGAGCTTCCCTCCAGCTTTCTCCAGCTTTCTGCTGAG 4619
Qy |||||
1621 AGGAACAGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1680
Db |||||
4620 AGGAACAGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 4679
Qy |||||
1681 TAACTGGGCAAGTGTCTTTTACCTGATTTGATGATACATTTTCAATTAAGGTTCCAGTTA 1740
Db |||||
4680 TAACTGGGCAAGTGTCTTTTACCTGATTTGATGATACATTTTCAATTAAGGTTCCAGTTA 4739
Qy |||||
1741 TAAATATTTTGTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1800
Db |||||
4740 TAAATATTTTGTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4799
Qy |||||
1801 TATTTTAAATAGCTAGTAAACATATGATGATATAATTTCTAGAAACAAACATCTAATA 1860
Db |||||
4800 TATTTTAAATAGCTAGTAAACATATGATGATATAATTTCTAGAAACAAACATCTAATA 4859
Qy |||||
1861 AGTATAATCTCTGAAATPATGAGGCTTGATATAATTTAGGTTGTCCAGATGAAGCATG 1920
Db |||||
4860 AGTATAATCTCTGAAATPATGAGGCTTGATATAATTTAGGTTGTCCAGATGAAGCATG 4919
Qy |||||
1921 CTAGAAGCTGTAAACAGATACATAGAGATTAATGAGGATTTATGATGAACCTTAATAT 1980
Db |||||
4920 CTAGAAGCTGTAAACAGATACATAGAGATTAATGAGGATTTATGATGAACCTTAATAT 4979
Qy |||||
1981 ATAATGTTGCCAGGATTTAGTTCAATATTTGTTACTGTTATCTATCTGCTGATATGG 2040
Db |||||
4980 ATAATGTTGCCAGGATTTAGTTCAATATTTGTTACTGTTATCTATCTGCTGATATGG 5039
Qy |||||
2041 AATTCCTTTTAAATTCAAACGCTGAAACGAATCAGCATTTAGTCTTTGCCAGGCACACCAA 2100

5040 AATTCCTTTTAAATTCAAACGCTGAAACGAATCAGCATTTAGTCTTTGCCAGGCACACCAA 5099
Db |||||
2101 TAATCAGTCATGTGTAATATATGACAAAGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTG 2160
Qy |||||
5100 TAATCAGTCATGTGTAATATATGACAAAGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTG 5159
Db |||||
2161 TTTTTCGTTTAAAGTTGATGATCTTCTGAGGAAATAGTCACTCATTCCCACTCCACA 2220
Qy |||||
5160 TTTTTCGTTTAAAGTTGATGATCTTCTGAGGAAATAGTCACTCATTCCCACTCCACA 5219
Db |||||
2221 TAAGGGGTTTATAGAGAGAGTCTCTGCTGCTGATGATGATGATGATGATGATGATGATGATG 2280
Qy |||||
5220 TAAGGGGTTTATAGAGAGAGTCTCTGCTGCTGATGATGATGATGATGATGATGATGATGATG 5279
Db |||||
2281 CCTTTCGTTTAAATAGTCACTCATTCTATGCGAAACAGGAAACGATCCATAACTTTAGT 2340
Qy |||||
5280 CCTTTCGTTTAAATAGTCACTCATTCTATGCGAAACAGGAAACGATCCATAACTTTAGT 5339
Db |||||
2341 CTTAATGACACATTTGCAATTTTGTATTAATTTTGTGTTTCTTTGAGGTTGATCG 2400
Qy |||||
5340 CTTAATGACACATTTGCAATTTTGTATTAATTTTGTGTTTCTTTGAGGTTGATCG 5399
Db |||||
2401 TTGCTGCTTTTGTGCTGCACTTTTACTTTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2460
Qy |||||
5400 TTGCTGCTTTTGTGCTGCACTTTTACTTTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5459
Db |||||
2461 GAAAGCTTTGGGATCTTCAATTAATTTAGTACGACTGCTCAACAGCGTGCAGGTTTCTGTTT 2520
Qy |||||
5460 GAAAGCTTTGGGATCTTCAATTAATTTAGTACGACTGCTCAACAGCGTGCAGGTTTCTGTTT 5519
Db |||||
2521 CTGCTGCTTTGGGATCTTCAATTAATTTAGTACGACTGCTCAACAGCGTGCAGGTTTCTGTTT 2580
Qy |||||
5520 CTGCTGCTTTGGGATCTTCAATTAATTTAGTACGACTGCTCAACAGCGTGCAGGTTTCTGTTT 5579
Db |||||
2581 TGTACCATATTTTGTGTAATTTTATGTTTCTTAAACAAATTTATCGTATAGGTTGA 2640
Qy |||||
5580 TGTACCATATTTTGTGTAATTTTATGTTTCTTAAACAAATTTATCGTATAGGTTGA 5639
Db |||||
2641 TGAACGCTCATGTGTTTGTCCAAAGAGCTGTAATATTTATTTATGTTTCACTGCTCAA 2700
Qy |||||
5640 TGAACGCTCATGTGTTTGTCCAAAGAGCTGTAATATTTATTTATGTTTCACTGCTCAA 5699
Db |||||
2701 AATTTCACTGACCTGAAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2760
Qy |||||
5700 AATTTCACTGACCTGAAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5759
Db |||||
2761 AATTAATTTGTAAAAAGGTTTCT 2785
Qy |||||
5760 AATTAATTTGTAAAAAGGTTTCT 5784
Db |||||

RESULT 10
ADE85079
ID ADE85079 standard; DNA; 5784 BP.
XX
AC ADE85079;
XX
DT 29-JAN-2004 (first entry)
XX
DE Farnesyl transferase inhibitor modulated leukemia associated gene #298.
XX ss; cytostatic; farnesyl transferase inhibitor; gene expression;
KW quininone; leukemia; cancer.
XX
OS Homo sapiens.
XX
PN WO2003038129-A2.
XX
PD 08-MAY-2003.
XX
PF 30-OCT-2002; 2002WO-US034784.
XX

QY 1681 TAACTGCGGACAGTTGCTTTTACCTGATTTGATGATACATTTCAATAAGGTTCCAGTTA 1740
Db TAACTGCGGACAGTTGCTTTTACCTGATTTGATGATACATTTCAATAAGGTTCCAGTTA 4739
QY 1741 TAAATATTTTGTATATTTATTTAGTGACTAGTAATGCAATTCATTTACAGTAAC 1800
Db TAAATATTTTGTATATTTATTTAGTGACTAGTAATGCAATTCATTTACAGTAAC 4799
QY 1801 TATTTTAAATATGCTTAGTAACACATATGTAGTATATTTCTAGAAACAAACATCTTAATA 1860
Db TATTTTAAATATGCTTAGTAACACATATGTAGTATATTTCTAGAAACAAACATCTTAATA 4859
QY 1861 AGTATATTAATCTGTGAAATAATGAGGCTTGATAATTTAGTGTGTGACGATGAAGCATG 1920
Db AGTATATTAATCTGTGAAATAATGAGGCTTGATAATTTAGTGTGTGACGATGAAGCATG 4919
QY 1921 CTAGAAGCTGTAAACAGATACATAGAGATATGAGGAGTTTATGATGGAACCTTAATAT 1980
Db CTAGAAGCTGTAAACAGATACATAGAGATATGAGGAGTTTATGATGGAACCTTAATAT 4979
QY 1981 ATAAATGTTGCCAGCGATTTTGTCTCAATATTTTACTGTATCTATCTGCTGATATGG 2040
Db ATAAATGTTGCCAGCGATTTTGTCTCAATATTTTACTGTATCTATCTGCTGATATGG 5039
QY 2041 AATTCCTTTAATCAACGCTGAAACGAAATCAGCATTTAGTCTTCCAGGCACACCCAA 2100
Db AATTCCTTTAATCAACGCTGAAACGAAATCAGCATTTAGTCTTCCAGGCACACCCAA 5099
QY 2101 TAATCAGTCATGTGTAATATGACAGTTGTTTTGTTTTGTTTTGTTTTGTTTTGTTTTG 2160
Db TAATCAGTCATGTGTAATATGACAGTTGTTTTGTTTTGTTTTGTTTTGTTTTGTTTTG 5159
QY 2161 TTTTTTCTTTAAGTTCATGATCTTCTGCGAGAAATAGTCATCTCCACTCCACCA 2220
Db TTTTTTCTTTAAGTTCATGATCTTCTGCGAGAAATAGTCATCTCCACTCCACCA 5219
QY 2221 TAAGGGTTTAGTAAGAGAGTCTGCTGCTGATGATGATGATGATGATGATGATGATG 2280
Db TAAGGGTTTAGTAAGAGAGTCTGCTGCTGATGATGATGATGATGATGATGATGATGATG 5279
QY 2281 CCCTTCTGTTAATAGTATCATCATTTCTATGCGCAACAGCAATCCATTAACCTTTAGT 2340
Db CCCTTCTGTTAATAGTATCATCATTTCTATGCGCAACAGCAATCCATTAACCTTTAGT 5339
QY 2341 CTTAATGTACATTCATTTGATAAATAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2400
Db CTTAATGTACATTCATTTGATAAATAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5399
QY 2401 TTGTGTTGTTTGTGCACTTTTACTTTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2460
Db TTGTGTTGTTTGTGCACTTTTACTTTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5459
QY 2461 GAACGGTTGGGATACCTCATTAATGTAGCACTGTCAACGCTGCAAGTTTCTGTTT 2520
Db GAACGGTTGGGATACCTCATTAATGTAGCACTGTCAACGCTGCAAGTTTCTGTTT 5519
QY 2521 CTGTGTTGTGGGGTCAACCGTACAAATGCTGCGGAATGACCATGATGTGAATATTTAGAA 2580
Db CTGTGTTGTGGGGTCAACCGTACAAATGCTGCGGAATGACCATGATGTGAATATTTAGAA 5579
QY 2581 TGTAACATATTTTGTGAAATATTTATGTTTTTCTAAACAAATTTATCGTATAGGTTGA 2640
Db TGTAACATATTTTGTGAAATATTTATGTTTTTCTAAACAAATTTATCGTATAGGTTGA 5639
QY 2641 TGAACGTGATGCTTTTGGCAAGACGTGAATATTTATTTATGTTTCAATGTTCAATGTTCAA 2700
Db TGAACGTGATGCTTTTGGCAAGACGTGAATATTTATTTATGTTTCAATGTTCAATGTTCAA 5699
QY 2701 AATTTCCACTGAAACCTGACCTTAGCTAGTAACCTCATTTTAAAGATTTAAACAGG 2760
Db AATTTCCACTGAAACCTGACCTTAGCTAGTAACCTCATTTTAAAGATTTAAACAGG 5759
QY 2761 AATAAATTTGAAAAAGGTTTTCT 2785

Db 5760 AATAAATTTGTAATAAAGGTTTTCT 5784

RESULT 11
ADQ18777

ID ADQ18777 standard; DNA; 5784 BP.

XX ADQ18777;

DT 26-AUG-2004 (first entry)

DE Human soft tissue sarcoma-upregulated DNA - SEQ ID 1596.

KW soft tissue sarcoma; cytostatic; gene therapy; vaccine; screening; human;
ds.

OS Homo sapiens.

XX WO2004048938-A2.

XX 10-JUN-2004.

XX 26-NOV-2003; 2003WO-US038193.

XX 26-NOV-2002; 2002US-0429739P.

XX (PROT-) PROTEIN DESIGN LABS INC.

XX Aziz N, Ginsburg WM, Zlotnik A;

XX WPI; 2004-441208/41.

XX Early detection of soft tissue sarcoma comprises determining expression
of a gene in a first soft tissue sample and a normal soft tissue sample
and comparing the gene expression, also useful in treating soft tissue
sarcoma.

XX Example 2; SEQ ID NO 1596; 210pp; English.

XX The invention relates to a novel method for detecting soft tissue sarcoma
which comprises obtaining a first soft tissue sample from an individual
and a normal soft tissue sample from the same or different individual,
determining the expression of a gene in both samples and comparing the
expression of the gene in both soft tissue samples, where a higher level
of protein expression in the first soft tissue sample indicates the
presence of soft tissue sarcoma. The method of the invention has
cytostatic applications and may be useful for detecting soft tissue
sarcoma, possibly via gene therapy or vaccine production. The nucleic
acid sequences may be useful in diagnostic and screening applications.
The current sequence is that of a human soft tissue sarcoma-upregulated
DNA of the invention. The current sequence is not shown within the
specification per se but was submitted in CD format by the inventor.

XX Sequence 5784 BP; 1447 A; 1460 C; 1518 G; 1359 T; 0 U; 0 Other;

Query Match 100.0%; Score 2785; DB 12; Length 5784;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2785; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAGGACTTGGACCGTGTATGACGGGGTGATATTTGTAAGATGATTTTGCAATGCAAC 60
Db 3000 GAGGACTTGGACCGTGTATGACGGGGTGATATTTGTAAGATGATTTTGCAATGCAAC 3059

QY 61 ATCCAGATATTTGATGTGTCTCTGAAACAATGCCATCAGTGACAGACTTCAGG 120

Db 3060 ATCCAGATATTTGATGTGTCTCTGAAACAATGCCATCAGTGACAGACTTCAGG 3119

QY 121 AACTTCAGATGTCCCTTTGGATCCCAAGGGACCAACCCAAATTTGATCCCAACTGGGTC 180

Db 3120 AACTTCAGATGTCCCTTTGGATCCCAAGGGACCAACCCAAATTTGATCCCAACTGGGTC 3179

QY 181 ATTGCCATCAAGCGAAGGAGCTGGTTTCAGACAGCCAACTCGGACCCCGCATCGTGTGA 240

Db	3180	ATTCGCCATCAAGGCAAGAGCTGGTTTCAGACAGCCAACTCGGACCCCGCATCGCTGTA	3239	Db	4260	AACAAGGCTCAATTCAATTTGAAGTAATGTTTTAGTAAGGAGAGATTAGAAGACAACAG	4319
Qy	241	GGTTTTGACAGATTGGGTTCTGGACTTCAGTGGACATTCCTACGTAAACACATGACCGG	300	Qy	1321	GCATAGCAATGACATPAAGCTACCGATTAACTAATCGGAACATGTAAACACAGTTACAAA	1380
Db	3240	GGTTTTGACAGATTGGGTTCTGGACTTCAGTGGACATTCCTACGTAAACACATGACCGG	3299	Db	4320	GCATAGCAATGACATPAAGCTACCGATTAACTAATCGGAACATGTAAACACAGTTACAAA	4379
Qy	301	GACGACGACTATGCTGGCTTCCTTTGGTTTACCAAGTCAAGCAGCCGCTTCATATGGTG	360	Qy	1381	ATAAACGAACTCTCCTCTTGTCTCAATGAAAGCCCTCATGTGACGTAGAGATGCAAGTT	1440
Db	3300	GACGACGACTATGCTGGCTTCCTTTGGTTTACCAAGTCAAGCAGCCGCTTCATATGGTG	3359	Db	4380	ATAAACGAACTCTCCTCTTGTCTCAATGAAAGCCCTCATGTGACGTAGAGATGCAAGTT	4439
Qy	361	ATGTGGAAGCAGGTGACGACGACCTACTGAGGAGGACAGCCCAAGCGGGCTTATGGCTAC	420	Qy	1441	TCATCAAGAAACAAAACATCTTCGAAATGGGTGTGACGCGGTTCCAGATGTGGATTTGGC	1500
Db	3360	ATGTGGAAGCAGGTGACGACGACCTACTGAGGAGGACAGCCCAAGCGGGCTTATGGCTAC	3419	Db	4440	TCATCAAGAAACAAAACATCTTCGAAATGGGTGTGACGCGGTTCCAGATGTGGATTTGGC	4499
Qy	421	TCGGGCTGTCCCTCAAGGTGGTGAATCCACACAGGGGACGGGACGACCTGAGGAAC	480	Qy	1501	AAAACTCTCAATTAAGTAAAGTTAGCAGAGCAAGTGCCTGCTTTAGTCTGCTTGT	1560
Db	3420	TCGGGCTGTCCCTCAAGGTGGTGAATCCACACAGGGGACGGGACGACCTGAGGAAC	3479	Db	4500	AAAACTCTCAATTAAGTAAAGTTAGCAGAGCAAGTGCCTGCTTTAGTCTGCTTGT	4559
Qy	481	GCGCTGTGGACACGCGGAAACACCGCGGGCAGGTGCGAACTTATGGCAGCAGCCCAAG	540	Qy	1561	GCGGTTGTGGGTCGGGAGGCTCCCTGCTGAGCTTCTTCCCGAGCTTTGCTGCTGAG	1620
Db	3480	GCGCTGTGGACACGCGGAAACACCGCGGGCAGGTGCGAACTTATGGCAGCAGCCCAAG	3539	Db	4560	GCGGTTGTGGGTCGGGAGGCTCCCTGCTGAGCTTCTTCCCGAGCTTTGCTGCTGAG	4619
Qy	541	AACATTTGGTGAAGGACTACAAGGCTTATAGTGGGACCTGACTCAGAGCCCAAGACC	600	Qy	1621	AGGAACGACAGCAGCAGCAGCGCGGAAAGGCGCATCTAAACGCGTATCTAGGCTTTGG	1680
Db	3540	AACATTTGGTGAAGGACTACAAGGCTTATAGTGGGACCTGACTCAGAGCCCAAGACC	3599	Db	4620	AGGAACGACAGCAGCAGCAGCGCGGAAAGGCGCATCTAAACGCGTATCTAGGCTTTGG	4679
Qy	601	GGCTACATCAGATCTTAGTCATGAAGGAAACAGGTCATGGCAGACTCAGGACCTATC	660	Qy	1681	TAACTGCGGCAAGTGTCTTTTACCTGATTTGATGATACATTTCAATTAAGTTCCAGTTA	1740
Db	3600	GGCTACATCAGATCTTAGTCATGAAGGAAACAGGTCATGGCAGACTCAGGACCTATC	3659	Db	4680	TAACTGCGGCAAGTGTCTTTTACCTGATTTGATGATACATTTCAATTAAGTTCCAGTTA	4739
Qy	661	TATGACCAAACTACGCTGGCGGGCTGGGTCTATTGTCTTCTCTCAAGAAATGGTC	720	Qy	1741	TAAATATTTTGTAAATATTTAATTAAGTGACATTAAGATGCAATCTCCATTTACCACTAACT	1800
Db	3660	TATGACCAAACTACGCTGGCGGGCTGGGTCTATTGTCTTCTCTCAAGAAATGGTC	3719	Db	4740	TAAATATTTTGTAAATATTTAATTAAGTGACATTAAGATGCAATCTCCATTTACCACTAACT	4799
Qy	721	TATTTCTCAGACCTCAAGTACGAATGACAGATATTTAAACAAGATTTGCTGCAATTCGG	780	Qy	1801	TATTTAAATAGCTAGTAAACATATGTAGTATAATTTCTAGAAAACAAACATCTAATA	1860
Db	3720	TATTTCTCAGACCTCAAGTACGAATGACAGATATTTAAACAAGATTTGCTGCAATTCGG	3779	Db	4800	TATTTAAATAGCTAGTAAACATATGTAGTATAATTTCTAGAAAACAAACATCTAATA	4859
Qy	781	GCAATGCCCTGTGATGTCATGCTCCCTAGACCTCAGTTCATTTGCTGCTCCGCT	840	Qy	1861	AGTATATACTCTGTGAAAATATGAGGCTTGTATATATTTAGTTGTGACGATGAAGCATG	1920
Db	3780	GCAATGCCCTGTGATGTCATGCTCCCTAGACCTCAGTTCATTTGCTGCTCCGCT	3839	Db	4860	AGTATATACTCTGTGAAAATATGAGGCTTGTATATATTTAGTTGTGACGATGAAGCATG	4919
Qy	841	TCTCTCTAGCAGCAGCTCCTCTGCTGCTTAACTTAACTGATGTTCTTCACTCCTG	900	Qy	1921	CTAGAGCTGTAAACAGATACATAGAGATAATGAGGAGTTTATGATGGAACCTTAAATAT	1980
Db	3840	TCTCTCTAGCAGCAGCTCCTCTGCTGCTTAACTTAACTGATGTTCTTCACTCCTG	3899	Db	4920	CTAGAGCTGTAAACAGATACATAGAGATAATGAGGAGTTTATGATGGAACCTTAAATAT	4979
Qy	901	CCAGCAACCCCAACCCCAAGTCCCTTACAGAGATAAATATCAATGGAATCTAGAGATGAA	960	Qy	1981	ATAATGTTGCCAGGAGTTTATGTTCAATATTTGTTACTTATCTATCTGCTGTATATGG	2040
Db	3900	CCAGCAACCCCAACCCCAAGTCCCTTACAGAGATAAATATCAATGGAATCTAGAGATGAA	3959	Db	4980	ATAATGTTGCCAGGAGTTTATGTTCAATATTTGTTACTTATCTATCTGCTGTATATGG	5039
Qy	961	CATCTAACCCACTAGAGGAACAGTTGGTGATATAGAGACTTTATGAGGAGTGAATA	1020	Qy	2041	AATCTTTTAAATCAACGCTGAAAACGAATCAGCATTTAGTCTTCCAGGACACCCCAA	2100
Db	3960	CATCTAACCCACTAGAGGAACAGTTGGTGATATAGAGACTTTATGAGGAGTGAATA	4019	Db	5040	AATCTTTTAAATCAACGCTGAAAACGAATCAGCATTTAGTCTTCCAGGACACCCCAA	5099
Qy	1021	TTGGGATGCCATTAATGCTTTTCTTTGTTTAAAGAAATGAGTTTACATATA	1080	Qy	2101	TAACTAGCTCATCTGTAAATATGACAAAGTTTGTGTTTGTGTTTGTGTTTGTGTTGG	2160
Db	4020	TTGGGATGCCATTAATGCTTTTCTTTGTTTAAAGAAATGAGTTTACATATA	4079	Db	5100	TAACTAGCTCATCTGTAAATATGACAAAGTTTGTGTTTGTGTTTGTGTTTGTGTTGG	5159
Qy	1081	AAATGTAATTAATATGTTATTTATGTTATGAGGTTGAAGGAAATCTGTGATAG	1140	Qy	2161	TTTTTTTGTCTTAAAGTTGCTGATGATCTTCTGAGGAAATAGTCACTCATCCCACTCCA	2220
Db	4080	AAATGTAATTAATATGTTATTTATGTTATGAGGTTGAAGGAAATCTGTGATAG	4139	Db	5160	TTTTTTTGTCTTAAAGTTGCTGATGATCTTCTGAGGAAATAGTCACTCATCCCACTCCA	5219
Qy	1141	CCATTTATGATAAATTAAGCATGAAATATGCTGAATCTTTTGGTCTTAAAGTTGT	1200	Qy	2221	TAAAGGTTTGTAGTAAGAGAACTCTGCTGTCTGATGATGATGAGGAGGAGGAGGAGG	2280
Db	4140	CCATTTATGATAAATTAAGCATGAAATATGCTGAATCTTTTGGTCTTAAAGTTGT	4199	Db	5220	TAAAGGTTTGTAGTAAGAGAACTCTGCTGTCTGATGATGATGAGGAGGAGGAGGAGG	5279
Qy	1201	CATATTCTTGAATAGAGTTGCTTCAATGACACACAAATCCCGCTAAATTAATATA	1260	Qy	2281	CCCTTTCTGTTAAATAGTCAATCAATTTCTATCCCAACAGGAAACGATCCATTAACCTT	2340
Db	4200	CATATTCTTGAATAGAGTTGCTTCAATGACACACAAATCCCGCTAAATTAATATA	4259	Db	5280	CCCTTTCTGTTAAATAGTCAATCAATTTCTATCCCAACAGGAAACGATCCATTAACCTT	5339
Qy	1261	AACAAGGCTCAATTCAAATTTGAAGTAATGTTTTAGTAAGGAGAGATTAGAACAACAG	1320	Qy	2341	CTTAATGTAACATTCGATTTTGAATAAATTAATTTTGTGTTTCTTTGAGGTTGATCG	2400
Db				Db	5340	CTTAATGTAACATTCGATTTTGAATAAATTAATTTTGTGTTTCTTTGAGGTTGATCG	5399

QY 901 CCAGCAACCCCAACCAAGTGCCTTCAGAGGATAAATATCAATGGAACCTCAGAGATGAA 960
Db 3910 CGAGCAACCCCAACCAAGTGCCTTCAGAGGATAAATATCAATGGAACCTCAGAGATGAA 3969
QY 961 CATCTAAACCCCACTAGAGGAAACCAAGTTGGTGATATATGAGACTTTATGTGGAGTGAAA 1020
Db 3970 CATCTAAACCCCACTAGAGGAAACCAAGTTGGTGATATATGAGACTTTATGTGGAGTGAAA 4029
QY 1021 TTGGGCAATGCCATACATACCTGCTTTCTGCTTTGTTTAAAGGAATACAGCTTTACATATA 1080
Db 4030 TTGGGCAATGCCATACATACCTGCTTTCTGCTTTGTTTAAAGGAATACAGCTTTACATATA 4089
QY 1081 AAATGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1140
Db 4090 AAATGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4149
QY 1141 CCATTATGATAAATTAAGCAATGAAATAATTAATTAATTAATTAATTAATTAATTAATTAAT 1200
Db 4150 CCATTATGATAAATTAAGCAATGAAATAATTAATTAATTAATTAATTAATTAATTAATTAAT 4209
QY 1201 CACTATTCTTGAATTAAGTGTCTCTACAATGACACAAATCCCGCTTAAATAAATTAATA 1260
Db 4210 CACTATTCTTGAATTAAGTGTCTCTACAATGACACAAATCCCGCTTAAATAAATTAATA 4269
QY 1261 AACAGGCTCAATTCAAATTTGAAGTAAATTTAGTAAGGAGAGATTTAGAAAGCAACAG 1320
Db 4270 AACAGGCTCAATTCAAATTTGAAGTAAATTTAGTAAGGAGAGATTTAGAAAGCAACAG 4329
QY 1321 GCATAGCAATTAAGCAATTAAGTAACTAATTCGAAACATGTAAGCAAGTTTACAAAA 1380
Db 4330 GCATAGCAATTAAGCAATTAAGTAACTAATTCGAAACATGTAAGCAAGTTTACAAAA 4389
QY 1381 ATAAACGAATCTCTCTCTGCTTCAATGAAGCCCTCATGTGAGTGAAGATGACAGTT 1440
Db 4390 ATAAACGAATCTCTCTCTGCTTCAATGAAGCCCTCATGTGAGTGAAGATGACAGTT 4449
QY 1441 TCATCAAGAAACAAACATCTTGCATTAAGTGGTGTGACGCGTTCCAGATGTGGATTTGGC 1500
Db 4450 TCATCAAGAAACAAACATCTTGCATTAAGTGGTGTGACGCGTTCCAGATGTGGATTTGGC 4509
QY 1501 AAAACCTCATTAAGTAAAGTTAGAGAGAGCAAGAGTGGTGTCTTTAGTGTCTGCTTGT 1560
Db 4510 AAAACCTCATTAAGTAAAGTTAGAGAGAGCAAGAGTGGTGTCTTTAGTGTCTGCTTGT 4569
QY 1561 GCGGTTGGCGTCCGGGAGGCTCCTGCTGAGTTTCTTCCCGAGCTTTGCTGCTGAG 1620
Db 4570 GCGGTTGGCGTCCGGGAGGCTCCTGCTGAGTTTCTTCCCGAGCTTTGCTGCTGAG 4629
QY 1621 AGGAACCCAGAGCAGCAGCAGCGCGGAAAGCGCATCTAACGCGTATCTAGGCTTTGG 1680
Db 4630 AGGAACCCAGAGCAGCAGCAGCGCGGAAAGCGCATCTAACGCGTATCTAGGCTTTGG 4689
QY 1681 TAACTGCGGACAAGTTGCTTTTACCTGATTTGATGATACATTTCAATTAAGGTTCCAGTTA 1740
Db 4690 TAACTGCGGACAAGTTGCTTTTACCTGATTTGATGATACATTTCAATTAAGGTTCCAGTTA 4749
QY 1741 TAAATATTTTGTATATATTTATTAAGTGAATAGCACTCAATTTACAGTAAC 1800
Db 4750 TAAATATTTTGTATATATTTATTAAGTGAATAGCACTCAATTTACAGTAAC 4809
QY 1801 TATTTTAAATATGCTTAGTAACATATGATATATTAATTTCTAGAAACAAACATCTAATA 1860
Db 4810 TATTTTAAATATGCTTAGTAACATATGATATATTAATTTCTAGAAACAAACATCTAATA 4869
QY 1861 AGTATATATCTGTGAAATATAGGCTTGTATATTAATTAAGTGTCTCAGATGAAGCATG 1920
Db 4870 AGTATATATCTGTGAAATATAGGCTTGTATATTAATTAAGTGTCTCAGATGAAGCATG 4929
QY 1921 CTAGAAGCTGTAAACAAATACATAGAGATTAATGAGGAGTTTATGATGAACTTAATAT 1980
Db 4930 CTAGAAGCTGTAAACAAATACATAGAGATTAATGAGGAGTTTATGATGAACTTAATAT 4989
QY 1981 ATAATGTTCCAGCGAATTTTGTGTTCAATATTTGTTTACTGTTATCTGCTGTATATGG 2040

Db 4990 ATAATGTTCCAGCGATTTTAGTTTCAATATTTGTTACTGTTATCTGCTGTATATGG 5049
QY 2041 AATTCCTTTAATTTCAAACGCTGAAAACGAATCAGCAATTTAGTCTTGCAGGACACCCAA 2100
Db 5050 AATTCCTTTAATTTCAAACGCTGAAAACGAATCAGCAATTTAGTCTTGCAGGACACCCAA 5109
QY 2101 TAATCAGTCAATGTATTAATGCAACAGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTG 2159
Db 5110 TAATCAGTCAATGTATTAATGCAACAGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTG 5169
QY 2160 ---GTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTG 2216
Db 5170 TTTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTG 5229
QY 2217 CACATAAGGGGTTTAGTAAGAGAGTCTGCTGCTGATGATGATGATGATGATGATGATGATGATG 2276
Db 5230 CACATAAGGGGTTTAGTAAGAGAGTCTGCTGCTGATGATGATGATGATGATGATGATGATGATG 5289
QY 2277 TTTCCCTTTCTGTTAATGATCATCATTTCTATGCAACAGAGAGATGATGATGATGATGATGATG 2336
Db 5290 TTTCCCTTTCTGTTAATGATCATCATTTCTATGCAACAGAGAGATGATGATGATGATGATGATGATG 5349
QY 2337 TAGCTTAATGATCAGATTCGATTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTG 2396
Db 5350 TAGCTTAATGATCAGATTCGATTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTG 5409
QY 2397 ATCCCTTG---TGTTGTTTGTGCTGCACTTTTACCTTTTGTGCTGAGCTGTATTCCTCG 2453
Db 5410 ATCCCTTGTTGTTGTTTGTGCTGCACTTTTACCTTTTGTGCTGAGCTGTATTCCTCG 5469
QY 2454 AGA-CACAGAGCGTTGGGATACCTTATTAATGATGAGGACTGTCAACAGCGTGCAGGTT 2512
Db 5470 AGACCAACGAGCGTTGGGATACCTTATTAATGATGAGGACTGTCAACAGCGTGCAGGTT 5529
QY 2513 TTCTGTTTCTGTTGTTGGGTCACCGTCAATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 2572
Db 5530 TTCTGTTTCTGTTGTTGGGTCACCGTCAATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 5589
QY 2573 ATTTAGATGATACCATATTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTG 2632
Db 5590 ATTTAGATGATACCATATTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTG 5649
QY 2633 TAGGTTGATGAAACGTCATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 2692
Db 5650 TAGGTTGATGAAACGTCATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 5709
QY 2693 ATGCTCAAAATTTCCACACTGAAACCTGACTAGTACCTGACTTTTAAAGATTG 2752
Db 5710 ATGCTCAAAATTTCCACACTGAAACCTGACTAGTACCTGACTTTTAAAGATTG 5769
QY 2753 ACACAGGAAATATAATTTGTTAAAGGTTTCT 2785
Db 5770 ACACAGGAAATATAATTTGTTAAAGGTTTCT 5802

RESULT 13

AAC77795

ID AAC77795 standard; cDNA; 3787 BP.

XX AAC77795;

AC AAC77795;

XX AAC77795;

DT 08-FEB-2001 (first entry)

XX Human cancer associated gene sequence SEQ ID NO:189.

DE Human; cancer associated gene; cancer antigen; detection; cancer;

XX diagnosis; cycostatic; proliferative; vulnary; immunomodulator;

KW antidiabetic; antihistaminic; antineoplastic; antineoplastic; antineoplastic;

KW antiinflammatory; antithyroid; antiallergic; antineoplastic; antineoplastic;

KW dermatological; neuroprotective; thrombolytic; coagulant; neurotropic;

KW vasotrophic; antipsoriatic; antiangiogenic; gene therapy; inflammation;

KW immune disorder; haematopoietic cell disorder; autoimmune disorder;

allergic reaction; graft versus host disease; organ rejection;
haemostatic; thrombolytic; cardiovascular disorder; infection;
neurological disease; drug screening; ss.

Homo sapiens.

WO20005350-A1.

21-SEP-2000.

08-MAR-2000; 2000WO-US005882.

12-MAR-1999; 99US-0124270P.

(HUMA-) HUMAN GENOME SCI INC.

Rosen CA, Ruben SM;

WPI; 2000-587533/55.

• SOCIETY •

never isolated nucleic acids comprising sequences encoding peptides useful for treating or diagnosing e.g. cancer.

Claim 1; Page 765-766; 2352pp; English.

RAC77607 to RAC78448 encode the human cancer associated proteins given in RAC43398 to AAB44239. The proteins can have activities based on the tissues and cells the genes are expressed in. Example of activities include: cytostatic; proliferative; vulnery; immunomodulator; antiidiabetic; antiasthmatic; antirheumatic; antiarthritic; antiinflammatory; antithyroid; aniallergic; antibacterial; antiviral; dermatological; neuroprotective; cardiac; thrombolytic; coagulant; nootropic; vasotropic; antipsoriatic and antiangiogenic. The polynucleotides and polypeptides can be used for preventing, treating or ameliorating medical conditions and diagnosing pathological conditions. Polynucleotides, polypeptides, antibodies, agonists and antagonists from the present invention may be used to treat immune disorders by activating or inhibiting the proliferation, differentiation or mobilisation of immune cells, to treat disorders of haematopoietic cells, autoimmune disorders, allergic reactions, graft versus host disease and organ rejection, modulate haemostatic or thrombolytic activity, modulate inflammation, cancers, cardiovascular disorders, neurological disease and bacterial or viral infections. The peptides, nucleotides, antibodies, agonists and antagonists may be also be used in drug screens. RAC78449 to RAC78457 and AAB44240 represent sequences used in the exemplification of the present invention

Sequence 3787 BP; 1049 A; 842 C; 893 G; 996 T; 0 U; 7 Other;

Query Match 98.2%; Score 2735.8; DB 3; Length 3787;
 Best Local Similarity 99.4%; Pred. NO. 0;
 Matches 2775; Conservative 4; Mismatches 6; Indels 8; Gaps 3;

Qy 1 GAGGAC TTGGACGGTGATGGACGGGGTGATATTTGTTAAAGATGATTTTGACAAATGCAAC 60
Db 983 GAGGAC TTGGACGGTGATGGACGGGGTGATATTTGTTAAAGATGATTTTGACAAATGCAAC 1042

QY 61 ATCCCGAGATATTGATGATGTGTCTCTGAAACAAATGCCATTCAGTGAGACAGACTTCAGG 120
Db 1043 ATCCCGAGATATTGATGATGTGTCTCTGAAACAAATGCCATTCAGTGAGACAGACTTCAGG 1102

Qy	121	AACTTC	CAGATG	CGTCC	CTTGAT	CCCAAA	GGGAC	CAACCA	AAATG	ATCGG	CAACT	GGGTC	180
Db	1103	AACTTC	CAGATG	CGTCC	CTTGAT	CCCAAA	GGGAC	CAACCA	AAATG	ATCGG	CAACT	GGGTC	

181 ATTCGCCATCAAGGCGCAGGAGCTGTTTCAGACAGCCAACTCGAACCCCGGCATCGCTGTA 240

DB	QY	Sequence	1222
1163	241	ATTCGGCATCAAGGCAAGGAGCTGGTTTCAGACAGCCAACTCGGACCCCGGCATCGCTGTA	1222

Qy	301	GACGACGACTATGCTGGCTTCGTTCTTTGGTTTACAGTCAAGCAGCCGCTCTCATGTGGTG	360
Db	1283	GACGACGACTATGCGGCTTCGTTCTTTGGTTTACAGTCAAGCAGCCGCTCTCATGTGGTG	1342
Qy	361	ATGTGGAAAGCAGGTGACGCAGACCTACTGGAGGACACAGCCCGACGCGGGCTATGGCTAC	420
Db	1343	ATGTGGAAAGCAGGTGACGCAGACCTACTGGAGGACACAGCCCGACGCGGGCTATGGCTAC	1402
Qy	421	TCCGGCGTGTCCCTCAAGSTGGTGAACTCCACACACGGGAGCGGCGAGCACCCTGAGGAAC	480
Db	1403	TCCGGCGTGTCCCTCAAGSTGGTGAACTCCACACACGGGAGCGGCGAGCACCCTGAGGAAC	1462
Qy	481	CGCTGTGGCACACGGGGAACACGCCGGGCGAGGTGCGAACTTTATGGCACGACCCACGG	540
Db	1463	CGCTGTGGCACASGGGGAAACACGCCGGGCGAGGTGCGAACCCTTATGGCACGACCCACGG	1522
Qy	541	AACATTGGCTGGAAGGACTACACGGCTATAGTGGGCACTGACTCACAGGCCCAAGACT	600
Db	1523	AACATTGGCTGGAAGGACTACACGGCTATAGTGGGCACTGACTCACAGGCCCAAGACT	1582
Qy	601	GGCTACATCAGAGTCTTAGTGCAATGAAGGAAAACAGTGCATGCGAGACTCAGGACCTATC	660
Db	1583	GGCTACATCAGAGTCTTAGTGCAATGAAGGAAAACAGTGCATGCGAGACTCAGGACCTATC	1642
Qy	661	TATGACCAAACTACGCTGGCGGGCGGCTGGGTCTATTGTGTTCTCTCAAGAAATGGTC	720
Db	1643	TATGACCAAACTACGCTGGCGGGCGGCTGGGTCTATTGTGTTCTCTCAAGAAATGGTC	1702
Qy	721	TATTTCTCAGACCTCAAGTAGGAAATGACAGATATTTAAACAAGATTGCTGCATTTCCG	780
Db	1703	TATTTCTCAGACCTCAAGTAGGAAATGACAGATATTTAAACAAGATTGCTGCATTTCCG	1762
Qy	781	GCAATGCCCTGTGCATGCCATGTTCCTAGACACCTCAGTTCATTCGTGGTCCCTGGGCT	840
Db	1763	GCAATGCCCTGTGCATGCCATGTTCCTAGACACCTCAGTTCATTCGTGGTCCCTGGGCT	1822
Qy	841	TCCTCTCTAGCAGCACCTCTGCTCCCTTGACCTTAACTCTGATGGTTCTTCACTCCCTG	900
Db	1823	TCCTCTCTAGCAGCACCTCTGCTCCCTTGACCTTAACTCTGATGGTTCTTCACTCCCTG	1882
Qy	901	CCAGCAACCCCAACCCAGTGCCTTCAGAGGATAAATATCAATGGAACTCAGAGATGAA	960
Db	1883	CCAGCAACCCCAACCCAGTGCCTTCAGAGGATAAATATCAATGGAACTCAGAGATGAA	1942
Qy	961	CATCTAACCCACTAGAGGAAACAGTTTGGTGATATGAGACTTTATGTGGGANGTGAANA	1020
Db	1943	CATCTAACCCACTAGAGGAAACAGTTTGGTGATATGAGACTTTATGTGGGANGTGAANA	2002
Qy	1021	TTGGGCATGCCATTACATTTGCTTTTGTGTTTAAAGAAATGACCTTTACATATA	1080
Db	2003	TTGGGCATGCCATTACATTTGCTTTTGTGTTTAAAGAAATGACCTTTACATATA	2062
Qy	1081	AAATGTAATTACTTATGTATTTATGTGTAATGAGTTGAAAGGAAATCTGTGCTAAG	1140
Db	2063	AAATGTAATTACTTATGTATTTATGTGTAATGAGTTGAAAGGAAATCTGTGCTAAG	2122
Qy	1141	CCATTATGATAAATTAAGCATCAAAAATTTGCTGAACTACTCTTTGGTGCCTTAAAGTTGT	1200
Db	2123	CCATTATGATAAATTAAGCATCAAAAATTTGCTGAACTACTCTTTGGTGCCTTAAAGTTGT	2182
Qy	1201	CACATTTCTTGAATTAGAGTTGCTCTCAATGACACACAAATCCCGCTAAAATAAATTATA	1260
Db	2183	CACATTTCTTGAATTAGAGTTGCTCTCAATGACACACAAATCCCGCTAAAATAAATTATA	2242
Qy	1261	AACAAGGTCATTTCAAATTTGAAAGTAATGTTTTAGTAAGGAGAGATTAGAGCAACAG	1320
Db	2243	AACAAGGTCATTTCAAATTTGAAAGTAATGTTTTAGTAAGGAGAGATTAGAGCAACAG	2302
Qy	1321	GCATAGCAAAATGACATAGCTTACCAATTAATTCGGAAACATGTAAACAGTTTACAAA	1380
Db	2303	GCATAGCAAAATGACATAGCTTACCAATTAATTCGGAAACATGTAAACAGTTTACAAA	2362

QY	1	GAGGACTTGGACGGGTGATGGACGGGGTGATATTTGTAAGATGATTTTGTGACAAATGACAAAC	60
Db	2892	GAGGACTTGGACGGGTGATGGACGGGGTGATATTTGTAAGATGATTTTGTGACAAATGACAAAC	2951
QY	61	ATCCACAGATATTGATGATCTGTGCTCTGAAACAATGCCATCAGTGAGACAGACTTCAGG	120
Db	2952	ATCCACAGATATTGATGATCTGTGCTCTGAAACAATGCCATCAGTGAGACAGACTTCAGG	3011
QY	121	AATTTCCAGATGCTCCCTTTGGATCCCAAGAGGACCCACCAAAATTGATCCCAACTGGGTC	180
Db	3012	AATTTCCAGATGCTCCCTTTGGATCCCAAGAGGACCCACCAAAATTGATCCCAACTGGGTC	3071
QY	181	ATTGCCATCAAGGCAAGAGCTGGTTCAGACAGCAAACTCGGACCCCGGCATCGCTGTA	240
Db	3072	ATTGCCATCAAGGCAAGAGCTGGTTCAGACAGCAAACTCGGACCCCGGCATCGCTGTA	3131
QY	241	GGTTTTGACGAGTTTGGGTCGTGGACTTCAGTGGCACAATTCIAGCTAAACACTGACCGG	300
Db	3132	GGTTTTGACGAGTTTGGGTCGTGGACTTCAGTGGCACAATTCIAGCTAAACACTGACCGG	3191
QY	301	GACGACGACTATGCTGGCTTCGTCTTTTGGTTACCAGTCAAGCAGCCGCTCTATGTGGTG	360
Db	3192	GACGACGACTATGCTGGCTTCGTCTTTTGGTTACCAGTCAAGCAGCCGCTCTATGTGGTG	3251
QY	361	ATGTGGAAGCAGGTGAACGACACTACTGTGGAGGACGAGCCACCGCGGGCCTATGGCTAC	420
Db	3252	ATGTGGAAGCAGGTGAACGACACTACTGTGGAGGACGAGCCACCGCGGGCCTATGGCTAC	3311
QY	421	TCCGGCGTCCCTCAAGTGGTGAACCTCCACACGCGGGGACGGCGGAGCAGCTTGAGGAAC	480
Db	3312	TCCGGCGTTCCTCTAAGTGGTGAACCTCCACACGCGGGGACGGCGGAGCAGCTTGAGGAAC	3371
QY	481	CGCTGTGGCACACGGGGGAAACACGCCGGGGCAGGTGCGAACTTATGGCACGACCCGAGG	540
Db	3372	CGCTGTGGCACACGGGGGAAACACGCCGGGGCAGGTGCGAACTTATGGCACGACCCGAGG	3431
QY	541	AACATTGGCTGGAAGGACTACAGCCCTATAGGTGGCACTCTGACTCACAGGCCCAAGACC	600
Db	3432	AACATTGGCTGGAAGGACTACAGCCCTATAGGTGGCACTCTGACTCACAGGCCCAAGACT	3491
QY	601	GGCTACATCAGAGCTTAGTGATGAAGGAAACAGSTCATGGGAGCTCAGGACCTATC	660
Db	3492	GGCTACATCAGAGCTTAGTGATGAAGGAAACAGGTCATGGCAGACTCAGGACCTATC	3551
QY	661	TATGACCAAACTACGCTGGCGGGCGCTGGGTCATTTCGCTCTCTCAAGAAATGGTC	720
Db	3552	TATGACCAAACTACGCTGGCGGGCGCTGGGTCATTTCGCTCTCTCAAGAAATGGTC	3611
QY	721	TATTTCTCAGACCTCAAGTACGAATGAGAGATATTTAAACAAGATTGCTGATTTCCG	780
Db	3612	TATTTCTCAGACCTCAAGTACGAATGAGAGATATTTAAACAAGATTGCTGATTTCCG	3671
QY	781	GCAATGCCCTGTGATGCCATGGTCCCTAGACACCTCAGTTCAITGTGTCCTTGGCGGT	840
Db	3672	GCAATGCCCTGTGATGCCATGGTCCCTAGACACCTCAGTTCAITGTGTCCTTGGCGGT	3731
QY	841	TCCTCTCTAGCAGCACCTCTGTCCTTCACTTAACTCTGATGCTTCTTCACTCCCTG	900
Db	3732	TCCTCTCTAGCAGCACCTCTGTCCTTCACTTAACTCTGATGCTTCTTCACTCCCTG	3791
QY	901	CCAGCAACCCCAACCCAGTGCCTTCAGAGGATAAATATCAATGAACTCAGAGATGAA	960
Db	3792	CCAGCAACCCCAACCCAGTGCCTTCAGAGGATAAATATCAATGAACTCAGAGATGAA	3851
QY	961	CATCTAACCCACTAGAGGAACAGTTTGGTGATATATGAGACTTTATGTGGAGTGAAA	1020
Db	3852	CATCTAACCCACTAGAGGAACAGTTTGGTGATATATGAGACTTTATGTGGAGTGAAA	3911
QY	1021	TTGGCATGCCATTACATTCCTTTTCTGTTTGGTTTAAAAAGAAATGAGCTTTACATATA	1080
Db	3912	TTGGCATGCCATTACATTCCTTTTCTGTTTGGTTTAAAAAGAAATGAGCTTTACATATA	3971
QY	1081	AAATGTAATTACTTATTTGATGTATATGGAGTTCAAGGGGAATTACTGTGCATAG	1140

Db	3972	AAATGTAATTA	CTTAATTTGTA	TTATTTGTA	TATGAGTTGA	AGGGAAT	PACTGTGCATAAG	4031
QY	1141	CCATTATGAT	AAATTAAGCAT	GACAAAAT	TATTCGTA	CACTACTTTT	TGGTCTTAAAGTTGT	1200
Db	4032	CCATTATGAT	AAATTAAGCAT	GACAAAAT	TATTCGTA	CACTACTTTT	TGGTCTTAAAGTTGT	4091
QY	1201	CACATACT	TTGAAATPAG	AGTTGCTCT	ACAATGAC	ACAAATCCG	CTAAATAAATTATA	1260
Db	4092	CACATACT	TTGAAATPAG	AGTTGCTCT	ACAATGAC	ACAAATCCG	CTAAATAAATTATA	4151
QY	1261	ACRAGGTC	CAATTC	CAAAATTTGA	GTAAATGTT	TAGTAAGG	AGATTAAGAACAACAG	1320
Db	4152	AACAAGGT	CAATTC	CAAAATTTGA	GTAAATGTT	TAGTAAGG	AGATTAAGAACAACAG	4211
QY	1321	GCATAGCA	AAATGAC	ATAAGCTAC	CGATTAACT	ATCGAATCG	AAATGACAAA	1380
Db	4212	GCATAGCA	AAATGAC	ATAAGCTAC	CGATTAACT	ATCGAATCG	AAATGACAAA	4271
QY	1381	ATAAACGA	ACTCTCCT	CTGTGCTCA	CAATGA	AAACCTCAT	GTGCGATGCGATT	1440
Db	4272	ATAAACGA	ACTCTCCT	CTGTGCTCA	CAATGA	AAACCTCAT	GTGCGATGCGATT	4331
QY	1441	TCATCAAG	ACAAACAT	CCCTTGCA	AAATGGG	TGTGACCG	GGTCCAGATGTGAATTGGC	1500
Db	4332	TCATCAAG	ACAAACAT	CCCTTGCA	AAATGGG	TGTGACCG	GGTCCAGATGTGAATTGGC	4391
QY	1501	AAAACCT	CATTAA	TAAAGTTAG	CAGACAA	AGTGGT	CTTACAGATGTGATTGGC	1560
Db	4392	AAAACCT	CATTAA	TAAAGTTAG	CAGACAA	AGTGGT	CTTACAGATGTGATTGGC	4451
QY	1561	GCGGTGTG	GGCTCGG	GAGGCTCT	CGCTGAG	CTTCCCG	CAGCTTTGCTGCGCTGAG	1620
Db	4452	GCGGTGTG	GGCTCGG	GAGGCTCT	CGCTGAG	CTTCCCG	CAGCTTTGCTGCGCTGAG	4511
QY	1621	AGGAACAG	ACAGAC	ACGACG	CGGAA	AGGCGCAT	CTAACCGTATCTAGGCTTTGG	1680
Db	4512	AGGAACAG	ACAGAC	ACGACG	CGGAA	AGGCGCAT	CTAACCGTATCTAGGCTTTGG	4571
QY	1681	TAACTCGG	ACAGCTG	CTTTACCT	GATTGAT	GATACATTT	CATTAAAGTTTCCAGTTA	1740
Db	4572	TAACTCGG	ACAGCTG	CTTTACCT	GATTGAT	GATACATTT	CATTAAAGTTTCCAGTTA	4631
QY	1741	TAAATATTT	TGTTAA	TATTTA	TAAAGT	GACTATAG	AATGCAACTCCATTTACCAGTAACT	1800
Db	4632	TAAATATTT	TGTTAA	TATTTA	TAAAGT	GACTATAG	AATGCAACTCCATTTACCAGTAACT	4691
QY	1801	TATTTTAA	ATATGCT	AGTAAC	ACATATG	ATATTAATTT	CTAGAAACAAACATCTAAFA	1860
Db	4692	TATTTTAA	ATATGCT	AGTAAC	ACATATG	ATATTAATTT	CTAGAAACAAACATCTAAFA	4751
QY	1861	AGTATAAT	ATCCTGTG	AAATATAG	CGCTGTA	ATAATTAG	GTGTGCAGATGAAGCATG	1920
Db	4752	AGTATAAT	ATCCTGTG	AAATATAG	CGCTGTA	ATAATTAG	GTGTGCAGATGAAGCATG	4811
QY	1921	CTAGAAG	CTGAAC	AGATACAT	AGAGATAT	ATAGGAG	GTTTATGATGGAACCTTAAATA	1979
Db	4812	CTAGAAG	CTGAAC	AGATACAT	AGAGATAT	ATAGGAG	GTTTATGATGGAACCTTAAATA	4871
QY	1980	TATAATG	TGCGAC	GGATTTAG	TTCAATAT	TGTTTACT	GTATCTATCTGCTGATATG	2039
Db	4872	TATAATG	TGCGAC	GGATTTAG	TTCAATAT	TGTTTACT	GTATCTATCTGCTGATATG	4931
QY	2040	GAATTCCT	TTTAAAT	CAACGCTG	AAACAAAT	CAGCATTTAG	CTTGCCAGGCACACCA	2099
Db	4932	GAATTCCT	TTTAAAT	CAACGCTG	AAACAAAT	CAGCATTTAG	CTTGCCAGGCACACCA	4990
QY	2100	ATAATCAG	TGTAATG	CAAGTTG	CTTTTGT	TTTTTTG	TTTTTTTGTGTTGGTTG	2159
Db	4991	ATAATCAG	TGTAATG	CAAGTTG	CTTTTGT	TTTTTTG	TTTTTTTGTGTTGGTTG	5050
QY	2160	-----	GTTTTTTG	TGCTTAA	GTGATG	ATTTCT	TGCGAGGAATATGATCACTCATCCCACT	2215

Db 5051 GTTTGTTTCTTTTAAAGTTCATGATCTTTCTGCGAGAAATAGTCACTCATCCACT 5110
QY 2216 CCACATAAGGGTTTGTAGTAAGAAAGTCTCTGCTGATGATGCGATAGGGGCAATCT 2275
Db 5111 CCACATAAGGGTTTGTAGTAAGAAAGTCTCTGCTGATGATGCGATAGGGGCAATCT 5170
QY 2276 TTTTCCCTTCTCTGTTAATAGTCATCATTCTCTATGCCAAACAGGAAACCATCAAACT 2335
Db 5171 TTTTCCCTTCTCTGTTAATAGTCATCATTCTCTATGCCAAACAGGAAACCATCAAACT 5230
QY 2336 TTAGTCTTAATGTACACATTTGCAATTTTGATAAAATTAATTTGTTGTTCTTTGAGTT 2395
Db 5231 TTAGTCTTAATGTACACATTTGCAATTTTGATAAAATTAATTTGTTGTTCTTTGAGTT 5290
QY 2396 GATCGTTG----TGTGTTTCTGTCACATTTTACTTTTTTTCGGTGTGAGCTGTATCCC 2452
Db 5291 GATCGTTGTTGTTGTTTCTGTCACATTTTACTTTTTTTCGGTGTGAGCTGTATCCC 5350
QY 2453 GAGA-CRACGAAGCGTTGGGATCTTCAATTAATGTAGCGACTGTCAACAGCGTGCAGGT 2511
Db 5351 GAGACCAACGAGCGTTGGGATCTTCAATTAATGTAGCGACTGTCAACAGCGTGCAGGT 5410
QY 2512 TTTCTGTTTCTGTTGTTGGGTCACCGTACAAATGTTGGGAAATGACGATGATGAA 2571
Db 5411 TTTCTGTTTCTGTTGTTGGGTCACCGTACAAATGTTGGGAAATGACGATGATGAA 5470
QY 2572 TATTAGAATGACATATTTTGTAAATTTATTTATGTTTCTTAAACAAATTTATCGT 2631
Db 5471 TATTAGAATGACATATTTTGTAAATTTATTTATGTTTCTTAAACAAATTTATCGT 5530
QY 2632 ATAGTTGATCAAAAGCTGATGTTTTCGCAAGAGCTGTAATATTTATTTATGTTTCA 2691
Db 5531 ATAGTTGATCAAAAGCTGATGTTTTCGCAAGAGCTGTAATATTTATTTATGTTTCA 5590
QY 2692 CATGTCAAAATTTTCAACACTGAAACCTGCACTTAGCTAGAACCTCATTTTAAAGATT 2751
Db 5591 CATGTCAAAATTTTCAACACTGAAACCTGCACTTAGCTAGAACCTCATTTTAAAGATT 5650
QY 2752 AACACAGGAATAAATTGTAAAGGTTTCT 2785
Db 5651 AACACAGGAATAAATTGTAAAGGTTTCT 5684

XX WPI; 2003-635698/60.
XX New genes regulated by MYCN activation, useful in gene therapy,
PT particularly for treating a subject with e.g. neuroblastoma or other
PT cancers, or for diagnosing, staging or monitoring the treatment of the
PT cancer.
XX
XX
PS Claim 1; SEQ ID NO 25; 27pp; English.
XX
CC This invention relates to novel isolated cDNAs that are differentially
CC expressed in MYCN activated cells. Specifically, it refers to
CC polynucleotide sequences that exhibit differential expression patterns in
CC cells activated by the transactivator MYCN, where MYCN is a proto-
CC oncogene that is amplified in neuroblastoma cells and is common in small
CC cell lung cancers. The present invention describes these cDNA molecules
CC as useful for in hybridization assays to detect expression of nucleic
CC acids (or complementary nucleic acids) in a present in a given sample, as
CC well as for screening assays by identifying molecules or compounds that
CC specifically bind the cDNA as a ligand and modulate function or activity.
CC Accordingly, these compositions exhibit cytostatic activity and can also
CC be used for gene therapy purposes. This polynucleotide sequence is a cDNA
CC that is differentially expressed in MYCN activated cells, given in an
CC exemplification of the invention. NOTE: This sequence does not appear in
CC the printed specification but has been obtained in electronic format from
CC the US Patent Office at
CC ftp.seqdata.uspto.gov/sequence.html?DocID=20030119009.
XX
SQ Sequence 5696 BP; 1438 A; 1424 C; 1483 G; 1351 T; 0 U; 0 Other;

Query Match 97.3%; Score 2709.6; DB 10; Length 5696;
Best Local Similarity 99.3%; Pred. No. 0;
Matches 2775; Conservative 0; Mismatches 9; Indels 10; Gaps 5;
QY 1 GAGCACTTGGACGCTGATGGACGGGTGATATTTGTAAGATGATTTTGACAATGACAAC 60
Db 2892 GAGCACTTGGACGCTGATGGACGGGTGATATTTGTAAGATGATTTTGACAATGACAAC 2951
QY 61 ATCCAGATATTGATGATGTTGTCCTGAAACAAATGTCATCAGTGACAGACTTCAGG 120
Db 2952 ATCCAGATATTGATGATGTTGTCCTGAAACAAATGTCATCAGTGACAGACTTCAGG 3011
QY 121 AACTTCCAGATGTTCCCTTGGATCCCAAGGGACCCCAAAATGATCCCAACTGGGTC 180
Db 3012 AACTTCCAGATGTTCCCTTGGATCCCAAGGGACCCCAAAATGATCCCAACTGGGTC 3071
QY 181 ATTGCCATCAAGCAAGGAGCTGTTTCAGACAGCAACTCGGACCCCGGATCGCTGA 240
Db 3072 ATTGCCATCAAGCAAGGAGCTGTTTCAGACAGCAACTCGGACCCCGGATCGCTGA 3131
QY 241 GGTTTTGACGAGTTTGGGTCGTGTTGACATTCAGTGGCACATTTAGTAAACACTGACCG 300
Db 3132 GGTTTTGACGAGTTTGGGTCGTGTTGACATTCAGTGGCACATTTAGTAAACACTGACCG 3191
QY 301 GACGACGATGCTGGCTTGGTCTTGTGTTACCAAGTCAAGCAGCGCTTCTATGTGGTG 360
Db 3192 GACGACGATGCTGGCTTGGTCTTGTGTTACCAAGTCAAGCAGCGCTTCTATGTGGTG 3251
QY 361 ATGTGGAAGCAGTGCACGACACCTTACTGGGAGGACCAAGCCCAAGCGGCGCTATGCTAC 420
Db 3252 ATGTGGAAGCAGTGCACGACACCTTACTGGGAGGACCAAGCCCAAGCGGCGCTATGCTAC 3311
QY 421 TCCGGCGTGTCCCTCAAGGTGGTGAATCCACACAGGGGACGGGCGAGCACTTGAGGAC 480
Db 3312 TCCGGCGTGTCCCTCAAGGTGGTGAATCCACACAGGGGACGGGCGAGCACTTGAGGAC 3371
QY 481 GCGCTGTGGCACACGGGGAAACACGCGGGGAGGTGCGAACTTATGGCAGCAGCCCGAG 540
Db 3372 GCGCTGTGGCACACGGGGAAACACGCGGGGAGGTGCGAACTTATGGCAGCAGCCCGAG 3431
QY 541 AACATTGGCTGGAAGGACTTACACGGCTATAGGTGGCACTGACTCAGAGGCCAGAGACC 600
Db 3432 AACATTGGCTGGAAGGACTTACACGGCTATAGGTGGCACTGACTCAGAGGCCAGAGACC 3491

RESULT 15
ADU56219
ID ADU56219 standard; cDNA; 5696 BP.
XX
XX
AC ADU56219;
XX
DT 06-MAY-2004 (first entry)
XX
DE Bovine cDNA differentially expressed in MYCN activated cells SeqID 25.
XX
KW bovine; differential expression; transactivator; proto-oncogene;
KW neuroblastoma; small cell lung cancer; cytostatic; gene therapy; ss;
KW MYCN activated cell.
XX
OS Bos taurus.
XX
XX US2003119009-A1.
XX
XX 26-JUN-2003.
XX
XX 25-FEB-2002; 2002US-00084817.
XX
XX 23-FEB-2001; 2001US-0270784P.
XX
XX (STUA/) STUART S G.
XX (NUCH/) NUCHTERN J G.
XX (PLON/) PLON S E.
XX (SHOH/) SHOHE J M.
XX
XX Stuart SG, Nuchtern JG, Plon SE, Shohet JM;
PI

QY 601 GGCTACATCAGAGTCTTAGTGCGATGAGAGAAAAACAGGTCATGGCAGACTCAGGACCTATC 660
Db 3492 GGCTACATCAGAGTCTTAGTGCGATGAGAGAAAAACAGGTCATGGCAGACTCAGGACCTATC 3551
QY 661 TATGACCAACCTACGCTGGCGGGCGCTGGGCTATTGGTCTTCTCAAGAAATGGTC 720
Db 3552 TATGACCAACCTACGCTGGCGGGCGCTGGGCTATTGGTCTTCTCAAGAAATGGTC 3611
QY 721 TATTTCTCAGACCTCAAGTACGAATGCAGAGATATTAAACAAGATTGTCGCAATTCGG 780
Db 3612 TATTTCTCAGACCTCAAGTACGAATGCAGAGATATTAAACAAGATTGTCGCAATTCGG 3671
QY 781 GCAATGCCCTGTCATGCCATGTCCTCCTAGACACCTCAGTTCATATGGTCTTGGCGGCT 840
Db 3672 GCAATGCCCTGTCATGCCATGTCCTCCTAGACACCTCAGTTCATATGGTCTTGGCGGCT 3731
QY 841 TCTCTCTCTAGCAGCACCTCTCTGCTCCCTTACCTTAACCTCTGATGGTCTTCACTCTCTG 900
Db 3732 TCTCTCTCTAGCAGCACCTCTCTGCTCCCTTACCTTAACCTCTGATGGTCTTCACTCTCTG 3791
QY 901 CCAGCAACCCCAACCAAGTGCCTTTCAGAGGATAAATATCAATGGAACCTCAGAGATGAA 960
Db 3792 CCAGCAACCCCAACCAAGTGCCTTTCAGAGGATAAATATCAATGGAACCGCAGATGAA 3851
QY 961 CATCTAACCCACTAGAGGAAACAGTTTGGTGATATPATCAGACTTTATGTGGAGTGAAAA 1020
Db 3852 CATCTAACCCACTAGAGGAAACAGTTTGGTGATATPATCAGACTTTATGTGGAGTGAAAA 3911
QY 1021 TTGGGATGCCATTAATCATTTGCTTTTCTGTTGTTTAAAGAAATGAGCTTTACATATA 1080
Db 3912 TTGGGATGCCATTAATCATTTGCTTTTCTGTTGTTTAAAGAAATGAGCTTTACATATA 3971
QY 1081 AAATGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1140
Db 3972 AAATGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4031
QY 1141 CCATTAATGATAAATTAAGCATGAAAAATATTTGCTGAACCTACTTTTGGTCTTAAAGTTGT 1200
Db 4032 CCATTAATGATAAATTAAGCATGAAAAATATTTGCTGAACCTACTTTTGGTCTTAAAGTTGT 4091
QY 1201 CACTATTCTTGAATTTAGAGTGTCTCTAATGACACACAAATCCCGCTTAAATTAATTA 1260
Db 4092 CACTATTCTTGAATTTAGAGTGTCTCTAATGACACACAAATCCCGCTTAAATTAATTA 4151
QY 1261 AACAGGCTCAATTCAAATTTGAAGTAATTTAGTAAGGAGATTTAGAGACACAG 1320
Db 4152 AACAGGCTCAATTCAAATTTGAAGTAATTTAGTAAGGAGATTTAGAGACACAG 4211
QY 1321 GCATAGCAATGACATAGCTACCGATTAACTAATTCGGAACATGTAAACAGTTTACAAA 1380
Db 4212 GCATAGCAATGACATAGCTACCGATTAACTAATTCGGAACATGTAAACAGTTTACAAA 4271
QY 1381 ATAAACGAATCTCTCTTGTCTACATGAAAGCCCTCATGTGCAGTAGAGATGCAAGTT 1440
Db 4272 ATAAACGAATCTCTCTTGTCTACATGAAAGCCCTCATGTGCAGTAGAGATGCAAGTT 4331
QY 1441 TCATCAAGAACAAACATCTCTTGAATGGGTGTGACGCGTTCACAGATGGGATTTGGC 1500
Db 4332 TCATCAAGAACAAACATCTCTTGAATGGGTGTGACGCGTTCACAGATGGGATTTGGC 4391
QY 1501 AAAACCTCATTTAAGTAAAGGTTAGCAGACCAAGTGGGTGCTTTAGCTGCTGCTTGT 1560
Db 4392 AAAACCTCATTTAAGTAAAGGTTAGCAGACCAAGTGGGTGCTTTAGCTGCTGCTTGT 4451
QY 1561 GCCGTGTGGCGTCGGGAGGCTCTCTGCTGAGCTTCTTCCCGAGCTTTGCTGCCTGAG 1620
Db 4452 GCCGTGTGGGTGTCGGGAGGCTCTCTGCTGAGCTTCTTCCCGAGCTTTGCTGCCTGAG 4511
QY 1621 AGGAACAGAGCAGACGACAGGCGGAAAGGCGCATCTAAAGCGTATCTAGGCTTTGG 1680
Db 4512 AGGAACAGAGCAGACGACAGGCGGAAAGGCGCATCTAAAGCGTATCTAGGCTTTGG 4571

QY 1681 TAACTGGGACAGTGTGCTTTTACCTGATTTGATGATACATTTCAATTAAGGTTCCAGTTA 1740
Db 4572 TAACTGGGACAGTGTGCTTTTACCTGATTTGATGATACATTTCAATTAAGGTTCCAGTTA 4631
QY 1741 TAAATATTTTGTAAATATTTAATTAAGTACTATAGAAATGCAATCCCATTTTACAGTAAC 1800
Db 4632 TAAATATTTTGTAAATATTTAATTAAGTACTATAGAAATGCAATCCCATTTTACAGTAAC 4691
QY 1801 TATTTAAATATGCTCTAGTAACACATATGTAGTATTAATTTCTAGAAACAAACATCTAATA 1860
Db 4692 TATTTAAATATGCTCTAGTAACACATATGTAGTATTAATTTCTAGAAACAAACATCTAATA 4751
QY 1861 AGTATATAATCTCTGABAAATATAGGCTTGTATTAATTTAGGTTGTCAGATGAAGCATG 1920
Db 4752 AGTATATAATCTCTGABAAATATAGGCTTGTATTAATTTAGGTTGTCAGATGAAGCATG 4811
QY 1921 CTAGAAGCTGTAAACAGAAATACATAGAGAAATATAGGAGCTTTATGATGAACCTT-AA 1979
Db 4812 CTAGAAGCTGTAAACAGAAATACATAGAGAAATATAGGAGCTTTATGATGAACCTTAAATA 4871
QY 1980 TATAATTTGCCAGCGATTTTATGTTCAATATTTGTTACTGTTATCTATCTGCTGTATATG 2039
Db 4872 TATAATTTGCCAGCGATTTTATGTTCAATATTTGTTACTGTTATCTATCTGCTGTATATG 4931
QY 2040 GAAATCTTTTAAATTCAAAGCGCTGAAACGAAATCAGCAATTTAGTCTTGCAGGACACCC 2099
Db 4932 GAAATCTTTTAAATTCAAAGCGCTGAAACGAAATCAGCAATTTAGTCTTGCAGGACACCC 4990
QY 2100 ATAATCAGTCTGTTAATATGCAACAAAGTTTGTGTTTGTGTTTGTGTTTGTGTTGTTG 2159
Db 4991 ATAATCAGTCTGTTAATATGCAACAAAGTTTGTGTTTGTGTTTGTGTTTGTGTTGTTG 5050
QY 2160 ---GTTTGTGTTTAAAGTTGCAATGATCTTCTGAGGAAATAGTCACTCATCCACT 2215
Db 5051 GTTGTGTTTGTGTTTAAAGTTGCAATGATCTTCTGAGGAAATAGTCACTCATCCACT 5110
QY 2216 CCACATAGGCGTTTGTAGTAAGAGAGTCTGCTGCTGATGATGATAGGCGGCAAAATCT 2275
Db 5111 CCACATAGGCGTTTGTAGTAAGAGAGTCTGCTGCTGATGATGATAGGCGGCAAAATCT 5170
QY 2276 TTTTCCCTCTCTGTTAATAGTCAATTTCTATGCAACAAAGGAGGAGGAGGAGGAGGAGG 2335
Db 5171 TTTTCCCTCTCTGTTAATAGTCAATTTCTATGCAACAAAGGAGGAGGAGGAGGAGGAGG 5230
QY 2336 TTAGCTTTAATGTCACATTCGATTTGATAAAAATTAATTTGTTGTTTCTTTCAGGTT 2395
Db 5231 TTAGCTTTAATGTCACATTCGATTTGATAAAAATTAATTTGTTGTTTCTTTCAGGTT 5290
QY 2396 GATCGTTG---TGTTGTTTGTGCTGACCTTTTACTTTTTCGCGTGGAGCTGATTTCCC 2452
Db 5291 GATCGTTGTTGTTGTTGCTGACCTTTTACTTTTCTGCGTGGAGCTGATTTCCC 5350
QY 2453 GAGA-CAACGAAGCGTTGGGATCTTCAATTAATGTTAGCGACTCTCAACAGCGTGCAGGT 2511
Db 5351 GAGACCAACGAGCGTTGGGATCTTCAATTAATGTTAGCGACTCTCAACAGCGTGCAGGT 5410
QY 2512 TTTCTGTTCTGTTGTTGGGTCACCGTACATGTTGGGAAATGACGATGATGAA 2571
Db 5411 TTTCTGTTCTGTTGTTGGGTCACCGTACATGTTGGGAAATGACGATGATGAA 5470
QY 2572 TATTTAGAAATGACCAATTTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 2631
Db 5471 TATTTAGAAATGACCAATTTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 5530
QY 2632 ATAGGTTGATGAAACGCTATGTTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 2691
Db 5531 ATAGGTTGATGAAACGCTATGTTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 5590
QY 2692 CATGGTCAAAATTTCCACCTGAAACCCCTGCACTTAGCTAGAACCTCATTTTAAAGATT 2751
Db 5591 CATGGTCAAAATTTCCACCTGAAACCCCTGCACTTAGCTAGAACCTCATTTTAAAGATT 5650
QY 2752 AACACAGGAAATTAATTTGTTAAAAAGGTTTCT 2785

Db 5651 AACAACGCGAATAAATTGTAATAAGGTTTCT 5684

Search completed: November 22, 2004, 17:22:57
Job time : 1248.75 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 22, 2004, 15:08:21 ; Search time 11525.2 Seconds
(without alignments)
11427.340 Million cell updates/sec

Title: US-10-017-724-1_COPY_3000_5784

Perfect score: 2785

Sequence: 1 gaggactggacgtgatggg.....aattgtaaaaagggtttct 2785

Scoring table: IDENTITY NUC

Gapop 10_0 , Gapext 1.0

Searched: 4526729 seqs, 23644849745 residues

Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb_ba.*

2: gb_htg.*

3: gb_in.*

4: gb_om.*

5: gb_ov.*

6: gb_pat.*

7: gb_ph.*

8: gb_pl.*

9: gb_pr.*

10: gb_ro.*

11: gb_sts.*

12: gb_sy.*

13: gb_un.*

14: gb_vl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	2785	100.0	5784	6	AX330488 Sequence
2	2785	100.0	5784	6	AX375587 Sequence
3	2785	100.0	5784	6	AX409513 Sequence
4	2785	100.0	5784	6	AX480837 Sequence
5	2785	100.0	5784	6	AX774982 Sequence
6	2785	100.0	5784	9	HUMTHRSPO
7	2703	97.1	3339	9	HS807201
8	2372.8	85.2	5505	11	BV177227
9	2372.8	85.2	5505	11	BV177983
10	2372.8	85.2	5505	11	BV178635
11	2372.8	85.2	5505	11	BV179851
12	1985	71.3	2026	11	G06722
13	1963.8	70.5	159175	9	BX322234
14	1694	60.8	2780	9	HUMTHRSPD
15	809	29.0	921	11	BV177085
16	809	29.0	921	11	BV177192
17	809	29.0	921	11	BV177656
18	793.8	28.5	2154	10	BC031843
19	783.8	28.1	4125	10	BC053702

20	767.2	27.5	4697	4	BTTHROM2
21	753.6	27.1	4108	6	BD129562
22	753.6	27.1	4108	10	MUSTSF2B
23	739.8	26.6	2764	6	CQ714089
24	687.2	24.7	814	11	BV166806
25	595.6	21.4	619	6	BD265321
26	595.6	21.4	619	6	AR401307
27	595.6	21.4	619	6	AX192760
28	594.8	21.4	848	6	BD216525
29	585.8	21.0	928	4	AF454919
30	540.2	19.4	637	6	AX261797
31	531.8	19.1	3537	5	CHKTHBSA
32	516	18.5	516	6	AX397018
33	497.6	17.9	561	11	BV167541
34	495.4	17.8	5269	4	AB005287
35	495.4	17.8	5289	6	BD129598
36	492.2	17.7	2162	4	BTRNAT1
37	491.8	17.7	7231	6	AR447356
38	491.8	17.7	7231	6	AX281758
39	490.2	17.6	4434	6	BD129560
40	490.2	17.6	4434	9	HSTHROMR
41	490.2	17.6	5719	6	CQ727671
42	490.2	17.6	5722	6	CQ776339
43	490.2	17.6	5722	6	AX094823
44	490.2	17.6	5722	6	AX411057
45	490.2	17.6	5722	6	AX587747

ALIGNMENTS

RESULT 1
AX330488
LOCUS AX330488
DEFINITION Sequence 997 from Patent WO0194629.
ACCESSION AX330488
VERSION AX330488.1 GI:18103466
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1
AUTHORS Young, P.E., Augustus, M., Carter, K.C., Ebner, R., Endress, G.,
Horrikan, S., Soppet, D.R. and Weaver, Z.
TITLE Cancer gene determination and therapeutic screening using signature
gene sets
JOURNAL Patent: WO 0194629-A 997 13-DEC-2001;
Avalon Pharmaceuticals (US)
FEATURES
Location/Qualifiers
source
1..5784
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

Query Match 100.0%; Score 2785; DB 6; Length 5784;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2785; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAGGACTTGGACGCGTATGACGGGGTGATATTTGTAAAGATGATTTTGACAATGACAAAC 60
DB 3000 GAGGACTTGGACGCGTATGACGGGGTGATATTTGTAAAGATGATTTTGACAATGACAAAC 3059

QY 61 ATCCAGATATTTGATGATGTGTCTCTGAAACAAATGCCATGACGACAGACTTCAGG 120
DB 3060 ATCCAGATATTTGATGATGTGTCTCTGAAACAAATGCCATGACGACAGACTTCAGG 3119

QY 121 AACTTCAGATGTCCTTGGATCCCAAGGGACCAACCAAAATTTGATCCCAACTGGGTC 180
DB 3120 AACTTCAGATGTCCTTGGATCCCAAGGGACCAACCAAAATTTGATCCCAACTGGGTC 3179

QY 181 ATTCGCCATCAGGACAGGAGCTGGTTTCACAGACCACTCGGACCCCGGATCGTGTA 240


```
QY 2401 TTGTTGTTTCTGACACTTTTACTTTTTCGTTGGAGCTGTATCCCGAGACAAC 2460
Db 5400 TTGTTGTTTCTGCTGCACTTTTACTTTTTCGTTGGAGCTGTATCCCGAGACAAC 5459
QY 2461 GAAGCGTTGGGATCTTCAATTAATGAGGAGCTGTCAACAGCGTGTTCCTGTTT 2520
Db 5460 GAAGCGTTGGGATCTTCAATTAATGAGGAGCTGTCAACAGCGTGTTCCTGTTT 5519
QY 2521 CTGTTGTTGGGCTCAACCGTACAATGTTGGGATGAGGAGTGAATATTTAGAA 2580
Db 5520 CTGTTGTTGGGCTCAACCGTACAATGTTGGGATGAGGAGTGAATATTTAGAA 5579
QY 2581 TGTACCATATTTTCTAAATATTTATGTTTCTTAAACAAATTTATCGTATAGTTGA 2640
Db 5580 TGTACCATATTTTCTAAATATTTATGTTTCTTAAACAAATTTATCGTATAGTTGA 5639
QY 2641 TGAAGGCTGATGTTTTCGCAAGAGCTGTAATATTTATGTTGTTCAATGTCAA 2700
Db 5640 TGAAGGCTGATGTTTTCGCAAGAGCTGTAATATTTATGTTGTTCAATGTCAA 5699
QY 2701 AATTTCAACACTGAAACCTGCACTTAGCTAGAACCTCATTTTAAAGATTAAACACG 2760
Db 5700 AATTTCAACACTGAAACCTGCACTTAGCTAGAACCTCATTTTAAAGATTAAACACG 5759
QY 2761 AATAAATGTAAGGTTTCT 2785
Db 5760 AATAAATGTAAGGTTTCT 5784

RESULT 2
AX375587
LOCUS AX375587
DEFINITION Sequence 3 from Patent WO0209735.
ACCESSION AX375587
VERSION AX375587.1 GI:19170155
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
Bornstein, P., Kyriakides, T., Ratner, B., Giachelli, C.,
Martinson, L.D. and Scatena, M.
Methods and devices to modulate the wound response
Patent: WO 0209735-A 3 07-FEB-2002;
The University of Washington (US)
FEATURES
Location/Qualifiers
1..5784
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
240..3758
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD26800.1"
/db_xref="GI:19170156"
/translation="MVMRLVLLALWVWPSTQAGHQDKDTPDLPESISNINRKTICAKQ
FRGPDGPVAVRFVFDYIPVNADDLSKITKIMRQKGFELTAQLQDKGSRGTLIA
LEGFLQSQRQREIVSNGPFDLDITMIDGTRHVVSLEDDGLADSQWKNVTVQAGET
YSLHVGCDLIGFVALDPFVLEHLQAEKSRMYVAKSARESHFRGLLQNLVLFENSVE
DILSKGQQQQQGAENAINISENTELRGLPHVTTEYVGPSSRPEVCEPCEBLGNM
VOELSGHLVNLQSENLRVSNQDNQFLWELIGGPPTKTRNSACWQDGRFAENETWV
VDSCTCTCKKFKFTIHOITCPATCASPSFVEGECPCSLHSVDGEGHSPAEWTO
CSVTGCGSTQORGBSDVTSNTCLGPSIQTRACLSKDCDTRIDGGGSHSPWSSCS
VTCGVGNITRILNCSVPVPMGKRNKSGRETKACQAPCIDGRSPWSPWACSV
TAGGIRERTKVCNPFQYGGKACVGDVQERQNCNKRSCVPDCLSNPCFPQAQCS
FPDGWSGCFPGFLNGTHCEDLDECALVPDI CFSTSKVPRCVNTPGPHCLPFP
RYRGNQPVGVEAKTEKQVCEPNCKDKTHKHAECIYLGHFSDPMYKCECOT
GYAGDGLICGSDLDLQWPNLNLVCATNATVHCIKDNCPLHNSGDEDFDKGLGDAC
DDDDNDGVTDKDNQCLLFPNRQADYDKDEVGDRNCNCPYVHPNPAQIDTDNNEGDA
CSVDIDGDDVFNERNDCPYVTNTQORDTDGVDGDCNDNCPVHPNPQTDVNDLVGD
QCDNNEIDDDGHQNNQDNCFYISNANQADHHRDQGDACDPPDDNDGVDPPDRNCLR
```

```
VENPOEDLDGDRGDIKDDFDNDONIPIIDDDVCENNNAISETDFRNFQWFLDPKGT
TQIDNNVIRHQGKELVQTPANSDPGIAVGDFSGVDFSGTFFVNTDRDDDDYAGVFG
YSSSRFYVMMKQVTQTYWEDQPTRAYGVSGLKVNSTTGTGSHLRNALMHTGNT
PGQVTLRHDPRNIGWKDYTAIRWHLTHRPKTYIRVLVHEGKQVMADSGPIYDQTYA
GGRLGLFVPSQBMVYFSDLKYECRDI"
ORIGIN
Query Match 100.0%; Score 2785; DB 6; Length 5784;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2785; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GAGGACTTCGACGGTGATGCGACGGGGTGATATTGTTAAAGATGATTTTGACAAATGACAAC 60
Db 3000 GAGGACTTCGACGGTGATGCGACGGGGTGATATTGTTAAAGATGATTTTGACAAATGACAAC 3059
QY 61 ATCCAGATATTGATGATGTTGCTCTGAAACAAATGCCATCAGTGAGACAGACTTCAGG 120
Db 3060 ATCCAGATATTGATGATGTTGCTCTGAAACAAATGCCATCAGTGAGACAGACTTCAGG 3119
QY 121 AACTTCCAGATGTTCCCTTGGATCCAAAGGACCAACCCAAATTTGATCCCACTGGGTC 180
Db 3120 AACTTCCAGATGTTCCCTTGGATCCAAAGGACCAACCCAAATTTGATCCCACTGGGTC 3179
QY 181 ATTCCGCATCAAGGCAAGGAGCTGTTTCAGACAGCAACTCGGACCCCGGCATCGCTGA 240
Db 3180 ATTCCGCATCAAGGCAAGGAGCTGTTTCAGACAGCAACTCGGACCCCGGCATCGCTGA 3239
QY 241 GGTTTGACAGTTTGGGTTCTGTGACATTCAGTGGCACATTTACGTAAACACTGACCCG 300
Db 3240 GGTTTGACAGTTTGGGTTCTGTGACATTCAGTGGCACATTTACGTAAACACTGACCCG 3299
QY 301 GACGAGCACTATGCTGGCTTCTGTTGGTTACCAAGTCAAGCAGCGCTTCTATGTTGGTG 360
Db 3300 GACGAGCACTATGCTGGCTTCTGTTGGTTACCAAGTCAAGCAGCGCTTCTATGTTGGTG 3359
QY 361 ATGTGGAAGCAGGTGACGACAGCACTTCTGGGAGGACCAAGCCACCGGGGCTTATGGCTAC 420
Db 3360 ATGTGGAAGCAGGTGACGACAGCACTTCTGGGAGGACCAAGCCACCGGGGCTTATGGCTAC 3419
QY 421 TCCGCGTGTCCCTCAAGGTGGTGAATCCACACGGGAGCGGAGCACTGAGGAAC 480
Db 3420 TCCGCGTGTCCCTCAAGGTGGTGAATCCACACGGGAGCGGAGCACTGAGGAAC 3479
QY 481 GCGCTGTGGCACAGGGGAAACAGCCGGGGGAGGTCGAACTTATGACAGACCCACG 540
Db 3480 GCGCTGTGGCACAGGGGAAACAGCCGGGGGAGGTCGAACTTATGACAGACCCACG 3539
QY 541 AACATTGGCTGGAAGGACTACACGCGCTTATAGTGGCACCTGACTCAGAGCCCAAGACC 600
Db 3540 AACATTGGCTGGAAGGACTACACGCGCTTATAGTGGCACCTGACTCAGAGCCCAAGACC 3599
QY 601 GGCTACATCAGAGTCTTAGTGCATGAAGGAAAACAGGTTCATGGCAGACTCAGGACTATC 660
Db 3600 GGCTACATCAGAGTCTTAGTGCATGAAGGAAAACAGGTTCATGGCAGACTCAGGACTATC 3659
QY 661 TATGACCAAACTACGCTGGCGGGGCTGGGTCTATTTTGTCTTCTCAAGAAATGTC 720
Db 3660 TATGACCAAACTACGCTGGCGGGGCTGGGTCTATTTTGTCTTCTCAAGAAATGTC 3719
QY 721 TATTTCTCAGACCTCAAGTACGAATGAGAGATATTTAAACAAGATTTGCTGCAATTTCCG 780
Db 3720 TATTTCTCAGACCTCAAGTACGAATGAGAGATATTTAAACAAGATTTGCTGCAATTTCCG 3779
QY 781 GCAATGCCCTGTGCATGCCATGGTCCCTAGACACTCAGTTTCTTGTGCTTCTTTCGGCT 840
Db 3780 GCAATGCCCTGTGCATGCCATGGTCCCTAGACACTCAGTTTCTTGTGCTTCTTTCGGCT 3839
QY 841 TCTCTCTTAGCAGCACTCTGTCCTCTGACCTTAACTCTGATGGTCTTCTTCACTCTG 900
Db 3840 TCTCTCTTAGCAGCACTCTGTCCTCTGACCTTAACTCTGATGGTCTTCTTCACTCTG 3899
QY 901 CCAGCAACCCCAACCAAGTGCCTTCCAGAGGATAAATATCAATGGAACCTCAGAGATGAA 960
Db 901 CCAGCAACCCCAACCAAGTGCCTTCCAGAGGATAAATATCAATGGAACCTCAGAGATGAA 960
```


3900 CCAGCAACCCCAACCCAGTGCCTTCAGAGGATATAATCAATGGAACCTCAGAGATGAA 3959
961 CATCTAACCCACTAGAGGAAACCAAGTTGGTGATATATGAGACTTTATGTGGAGTGAAGA 1020
3960 CATCTAACCCACTAGAGGAAACCAAGTTGGTGATATATGAGACTTTATGTGGAGTGAAGA 4019
1021 TTGGGCATGCAATTCACATTCGCTTTCTGTTGCTTTTAAAGAAATGACGCTTTACATATA 1080
4020 TTGGGCATGCAATTCACATTCGCTTTCTGTTGCTTTTAAAGAAATGACGCTTTACATATA 4079
1081 AAATGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1140
4080 AAATGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4139
1141 CCATTATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1200
4140 CCATTATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4199
1201 CACTATTCTTGAATAGAGTTGCTCTACAATGACACACAAATCCCGCTAAATAAATTAATA 1260
4200 CACTATTCTTGAATAGAGTTGCTCTACAATGACACACAAATCCCGCTAAATAAATTAATA 4259
1261 AACAGGCTCAATTCAAATTTGAAGTAATGTTTAAAGGAGAGATTTAGAGACACAG 1320
4260 AACAGGCTCAATTCAAATTTGAAGTAATGTTTAAAGGAGAGATTTAGAGACACAG 4319
1321 GCATAGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1380
4320 GCATAGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4379
1381 ATAAAGCAATCTCTCTGTTGCTTCAATTAATGAAAGCCCTCATGTGAGGAGATGCAAGTT 1440
4380 ATAAAGCAATCTCTCTGTTGCTTCAATTAATGAAAGCCCTCATGTGAGGAGATGCAAGTT 4439
1441 TCATCAAGAAACAAACATCTTGTGCAATGAAAGCCCTCATGTGAGGAGATGCAAGTTGCG 1500
4440 TCATCAAGAAACAAACATCTTGTGCAATGAAAGCCCTCATGTGAGGAGATGCAAGTTGCG 4499
1501 AAAACCTCATTTAAGTAAGTTAGCAGACCAAGTGGGTCCTTTAGCTGCTGCTGTT 1560
4500 AAAACCTCATTTAAGTAAGTTAGCAGACCAAGTGGGTCCTTTAGCTGCTGCTGTT 4559
1561 GCGGTTGTGGGTCGCGGAGGCTCTGCTGCTGAGCTTCCCTCCCAAGCTTTGCTGCTGAG 1620
4560 GCGGTTGTGGGTCGCGGAGGCTCTGCTGCTGAGCTTCCCTCCCAAGCTTTGCTGCTGAG 4619
1621 AGGAACCAAGACAGCAGACAGGCGGAAAGGCGCATCTAACGCGTATCTAGGCTTTGG 1680
4620 AGGAACCAAGACAGCAGACAGGCGGAAAGGCGCATCTAACGCGTATCTAGGCTTTGG 4679
1681 TAACTGCGGACAGGTTGCTTTTACCTGATTTGATGATACATTTCAATTAAGGTTCCAGTTA 1740
4680 TAACTGCGGACAGGTTGCTTTTACCTGATTTGATGATACATTTCAATTAAGGTTCCAGTTA 4739
1741 TAAATATTTGTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1800
4740 TAAATATTTGTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4799
1801 TATTTTAAATATGCTTAGTAACACATATGATATTAATTTCTAGAAAACAAATCTAATA 1860
4800 TATTTTAAATATGCTTAGTAACACATATGATATTAATTTCTAGAAAACAAATCTAATA 4859
1861 AGTATATATCTGTGAAATATGAGGCTTGATATTAATTAAGTTGTCAGATGAAGCATG 1920
4860 AGTATATATCTGTGAAATATGAGGCTTGATATTAATTAAGTTGTCAGATGAAGCATG 4919
1921 CTAGAAAGCTGTAAACAGATACATAGAGAAATATGAGGAGTTTATGATGAAACCTTAATAT 1980
4920 CTAGAAAGCTGTAAACAGATACATAGAGAAATATGAGGAGTTTATGATGAAACCTTAATAT 4979
1981 ATAAATGTCGACGAGTTTATGATTAATTTGTTTACTGTTATCTATCTGCTGATATGG 2040
4980 ATAAATGTCGACGAGTTTATGATTAATTTGTTTACTGTTATCTATCTGCTGATATGG 5039

2041 AATTCTTTTAATTCAAACGCTGAAACGAATCAGCAATTTAGTCTTGCAGGACACCCAA 2100
5040 AATTCTTTTAATTCAAACGCTGAAACGAATCAGCAATTTAGTCTTGCAGGACACCCAA 5099
2101 TAATCAGTCAATGATTAATGCAACAGTTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTG 2160
5100 TAATCAGTCAATGATTAATGCAACAGTTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTG 5159
2161 TTTTGTGTTTAAAGTTGATGATCTTCTGAGGAAATAGTCACTCATCCACTCCACA 2220
5160 TTTTGTGTTTAAAGTTGATGATCTTCTGAGGAAATAGTCACTCATCCACTCCACA 5219
2221 TAAGGGGTTTAAAGGAAAGTCTGCTGCTGATGATGATGATGATGATGATGATGATGATG 2280
5220 TAAGGGGTTTAAAGGAAAGTCTGCTGCTGATGATGATGATGATGATGATGATGATGATG 5279
2281 CCCTTCTGTTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2340
5280 CCCTTCTGTTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 5339
2341 CTTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2400
5340 CTTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 5399
2401 TTGTGTTGTTTGTGCTGACATTTTACATTTTGTGCTGCTGCTGCTGCTGCTGCTGCTG 2460
5400 TTGTGTTGTTTGTGCTGACATTTTACATTTTGTGCTGCTGCTGCTGCTGCTGCTGCTG 5459
2461 GAAGCGTTGGGATCTTCAATTAATGATGATGATGATGATGATGATGATGATGATGATG 2520
5460 GAAGCGTTGGGATCTTCAATTAATGATGATGATGATGATGATGATGATGATGATGATG 5519
2521 CTGTGTTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2580
5520 CTGTGTTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5579
2581 TGTACCATTAATTTTGTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2640
5580 TGTACCATTAATTTTGTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 5639
2641 TGAACGCTCATGCTGTTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2700
5640 TGAACGCTCATGCTGTTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5699
2701 AATTTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2760
5700 AATTTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5759
2761 AAATAAATTTGTAATAAAGGTTTCT 2785
5760 AAATAAATTTGTAATAAAGGTTTCT 5784

RESULT 3

AX409513

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Homo sapiens

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1

Alvares, C., Horne, D., Peres-da-Silva, S. and Vockley, J. G.

Gene expression profiles in liver cancer

Patent: WO 0229103-A 2160 11-APR-2002;

GENE LOGIC INC (US)

FEATURES

Location/Qualifiers

1..5784

/organism="Homo sapiens"

source

AX409513
Sequence 2160 from Patent WO0229103.
linear PAT 14-JUN-2002

AX409513

AX409513.1

GI:21442218

Homo sapiens (human)

Homo sapiens

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1

Alvares, C., Horne, D., Peres-da-Silva, S. and Vockley, J. G.

Gene expression profiles in liver cancer

Patent: WO 0229103-A 2160 11-APR-2002;

GENE LOGIC INC (US)

FEATURES

Location/Qualifiers

1..5784

/organism="Homo sapiens"

source

ORIGIN		/mol_type="unassigned DNA" /db_xref="taxon:9606" /note="EMBL/GenBank Accession No. L12350"	
Query Match		100.0%; Score 2785; DB 6; Length 5784;	
Best Local Similarity		100.0%; Pred. No. 0;	
Matches 2785; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
QY	1	GAGGACTTGGACCGGTGATGACGGGGTGATTTTGAAGATCATTTTGACATGCAAC	60
DB	3000	GAGGACTTGGACCGGTGATGACGGGGTGATTTTGAAGATCATTTTGACATGCAAC	3059
QY	61	ATCCAGATATTGATGATGTGTCTCTGAAAAAACAATGCCATCATGAGACAGACTTCAGG	120
DB	3060	ATCCAGATATTGATGATGTGTCTCTGAAAAAACAATGCCATCATGAGACAGACTTCAGG	3119
QY	121	AACCTCCAGATGTCCTCCCTTGGATCCCAAGGACCAACCAATGATCCCACTGGGTC	180
DB	3120	AACCTCCAGATGTCCTCCCTTGGATCCCAAGGACCAACCAATGATCCCACTGGGTC	3179
QY	181	ATTTCGCATCAAGGCAAGAGCTGTGTTACAGACAGCAACTCCGACCCCGCATCGCTGA	240
DB	3180	ATTTCGCATCAAGGCAAGAGCTGTGTTACAGACAGCAACTCCGACCCCGCATCGCTGA	3239
QY	241	GGTTTGGACGATTGGGTCTGTGGACTTCAGTGGCACATTCTACGTAAACACTGACCGG	300
DB	3240	GGTTTGGACGATTGGGTCTGTGGACTTCAGTGGCACATTCTACGTAAACACTGACCGG	3299
QY	301	GACGACGATGCTGGCTTCCTTTGGTTACAGTCAAGCAGCGCTTCTATGTGGTG	360
DB	3300	GACGACGATGCTGGCTTCCTTTGGTTACAGTCAAGCAGCGCTTCTATGTGGTG	3359
QY	361	ATGTGGAAAGCAGTCAAGCAGACTTACTGGGAGGACCAACCGGCGAGCACTTGAGGAA	420
DB	3360	ATGTGGAAAGCAGTCAAGCAGACTTACTGGGAGGACCAACCGGCGGCTTATGGCTAC	3419
QY	421	TCCGCGGTTCCTCAAGGTGTGAATCTCACACGGGACCGGCGAGCACTTGAGGAA	480
DB	3420	TCCGCGGTTCCTCAAGGTGTGAATCTCACACGGGACCGGCGAGCACTTGAGGAA	3479
QY	481	GGCTGTGGCAACCGGGNACACGGCGGGGAGGTGCGAACTTATGGACACCCACG	540
DB	3480	GGCTGTGGCAACCGGGNACACGGCGGGGAGGTGCGAACTTATGGACACCCACG	3539
QY	541	AACATTGGCTGAAGGACTACACGGCTTATAGTGGCACTTACCTCAAGGACCAAGAC	600
DB	3540	AACATTGGCTGAAGGACTACACGGCTTATAGTGGCACTTACCTCAAGGACCAAGAC	3599
QY	601	GGCTACATCAGAGTCTTAGTGATGAAGGAAACAGGTCTAGGCACTTACGAGCTTATC	660
DB	3600	GGCTACATCAGAGTCTTAGTGATGAAGGAAACAGGTCTAGGCACTTACGAGCTTATC	3659
QY	661	TATGACCAAACTAGCTGGGGGGGCTGGGTCTATTTGGTCTCTCAAGAAATGTC	720
DB	3660	TATGACCAAACTAGCTGGGGGGGCTGGGTCTATTTGGTCTCTCAAGAAATGTC	3719
QY	721	TATTTCTCAGACCTCAAGTACGAATGACAGATATTTAAACAAGATTTGCTGCAATTCG	780
DB	3720	TATTTCTCAGACCTCAAGTACGAATGACAGATATTTAAACAAGATTTGCTGCAATTCG	3779
QY	781	GCAATGCCCTGTGCATGCCATGGTCCCTAGACACCTCAGTTTCATTTGGTCTCTGGGCT	840
DB	3780	GCAATGCCCTGTGCATGCCATGGTCCCTAGACACCTCAGTTTCATTTGGTCTCTGGGCT	3839
QY	841	TCTCTCTAGCAGCACTCTCTGCTCCCTTGAACCTTAACTCTGATGGTTCCTCAGCTCTG	900
DB	3840	TCTCTCTAGCAGCACTCTCTGCTCCCTTGAACCTTAACTCTGATGGTTCCTCAGCTCTG	3899
QY	901	CCAGCAACCCCAACCCCAAGTGCCTTCAGAGGATAAATATCAATGGAACTCAGAGTAA	960
DB	3900	CCAGCAACCCCAACCCCAAGTGCCTTCAGAGGATAAATATCAATGGAACTCAGAGTAA	3959

QY	961	CATCTAACCCACTAGAGGAAACCCAGTTTGGTGATATATGAGACTTTATGTGGAGTGAAAA	1020
DB	3960	CATCTAACCCACTAGAGGAAACCCAGTTTGGTGATATATGAGACTTTATGTGGAGTGAAAA	4019
QY	1021	TTGGGCATGCCAATTACATTCGCTTTTCTTGTGTGTTTAAAAAGAAATGACGTTTACATATA	1080
DB	4020	TTGGGCATGCCAATTACATTCGCTTTTCTTGTGTGTTTAAAAAGAAATGACGTTTACATATA	4079
QY	1081	AAATGTAAATTAATTAATGTATTTATGTGTATATGAGGTTGAAAGGAAATCTGTGCATAAG	1140
DB	4080	AAATGTAAATTAATTAATGTATTTATGTATATGAGGTTGAAAGGAAATCTGTGCATAAG	4139
QY	1141	CAATTATGATAAATTAAGCATGAATAATCTGTAACACTTTTGGTGTCTTAAAGTTGT	1200
DB	4140	CAATTATGATAAATTAAGCATGAATAATCTGTAACACTTTTGGTGTCTTAAAGTTGT	4199
QY	1201	CACATTTCTTGAATTTAGAGTTGCTCTACAAATGACACACAAATCCCGCTAAATAATATA	1260
DB	4200	CACATTTCTTGAATTTAGAGTTGCTCTACAAATGACACACAAATCCCGCTAAATAATATA	4259
QY	1261	AACAGGGTCAATTCMAATTTGAAGTAATGTTTATAGTAGAGAGATTAGAGACAAACAG	1320
DB	4260	AACAGGGTCAATTCMAATTTGAAGTAATGTTTATAGTAGAGAGATTAGAGACAAACAG	4319
QY	1321	GCATAGCAAAATGACATAAGCTACCGATTAACTAAATCGGAACATGTAAACACAGTTACA	1380
DB	4320	GCATAGCAAAATGACATAAGCTACCGATTAACTAAATCGGAACATGTAAACACAGTTACA	4379
QY	1381	ATPAAACGAACCTCTCTTGTCTTCAATGAAAGCCCTCATGTGAGTAGAGATGAGATG	1440
DB	4380	ATPAAACGAACCTCTCTTGTCTTCAATGAAAGCCCTCATGTGAGTAGAGATGAGATG	4439
QY	1441	TCATCAAAAGAAACAAATCTTGGTGAATGGGTGTGACGGGTTCCAGATGTGATTTGGC	1500
DB	4440	TCATCAAAAGAAACAAATCTTGGTGAATGGGTGTGACGGGTTCCAGATGTGATTTGGC	4499
QY	1501	AAAACTCATTTAAGTAAAGGTTAGCAGAGCAAAAGTGGGTGCTTAGCTGCTGCTTGT	1560
DB	4500	AAAACTCATTTAAGTAAAGGTTAGCAGAGCAAAAGTGGGTGCTTAGCTGCTGCTTGT	4559
QY	1561	GCGGTGTGGGCTCGGGGAGGCTCTGCTGCTGAGCTTCTTCCCGAGCTTTGCTGCTGAG	1620
DB	4560	GCGGTGTGGGCTCGGGGAGGCTCTGCTGCTGAGCTTCTTCCCGAGCTTTGCTGCTGAG	4619
QY	1621	AGGAACACAGAGCAGCAGCAGCGGCGGAAAGGCGCATCTAACGGGTATCTAGGCTTTGG	1680
DB	4620	AGGAACACAGAGCAGCAGCAGCGGCGGAAAGGCGCATCTAACGGGTATCTAGGCTTTGG	4679
QY	1681	TAACTGCGGACAAAGTTGCTTTTACCTGATTTGATGATACATTTCAATTAAGGTTCCAGTTA	1740
DB	4680	TAACTGCGGACAAAGTTGCTTTTACCTGATTTGATGATACATTTCAATTAAGGTTCCAGTTA	4739
QY	1741	TAAATATTTTGTATTAATTTAATTAAGTACATATAGAAATGCAACTCCATTTACCAGTAAT	1800
DB	4740	TAAATATTTTGTATTAATTTAATTAAGTACATATAGAAATGCAACTCCATTTACCAGTAAT	4799
QY	1801	TATTTAAATATGCTAGTAAACACATATGATATAATTTCTAGAAACAAACATCTAATA	1860
DB	4800	TATTTAAATATGCTAGTAAACACATATGATATAATTTCTAGAAACAAACATCTAATA	4859
QY	1861	AGTATATACTCTGGAATATGAGGCTTGATATAATTAGGTTGTACGATGAAGCATG	1920
DB	4860	AGTATATACTCTGGAATATGAGGCTTGATATAATTAGGTTGTACGATGAAGCATG	4919
QY	1921	CTAGAAGCTGTAAACAGAAATACATAGAGAAATAAGGAGTTTATGATGGACCTTAATAT	1980
DB	4920	CTAGAAGCTGTAAACAGAAATACATAGAGAAATAAGGAGTTTATGATGGACCTTAATAT	4979
QY	1981	ATAATGTTCAGCGATTTTAGTTCAATTAATTTGTTACTGTATCTATCTGCTGTATATGG	2040
DB	4980	ATAATGTTCAGCGATTTTAGTTCAATTAATTTGTTACTGTATCTATCTGCTGTATATGG	5039
QY	2041	AAATCTTTTAAATCAACGGCTGAAACGAATCAGATTTTAGTCTTTGCCAGCACCCCAA	2100

3960	Db	CATCTAACCCCACTAGAGAAACCAAGTTTGGTGATATATAGACATTTTATGTGGAGTGA AAAA	4010
1021	QY	TTGGGCATGCCATTACATGTGCTTTTCTTGTTTCTTTAAAAAGAAATCACGTTTACATATA	1080
4020	Db	TTGGGCATGCCATTACATGTGCTTTTCTTGTTTAAAAAGATGACGTTTACATATA	4079
1081	QY	AAATGTAAATTACTTATTGTATTATGTATATGAGTTTGAAGGGAATATCTGTGCATAAG	1140
4080	Db	AAATGTAAATTACTTATTGTATTTATGTGTATATGAGTTTGAAGGGAATATCTGTGCATAAG	4139
1141	QY	CCATTATGATAAAATTAAGCATGAAAAATATTGCTGCACTACTTTTGGTGCTTAAAGTTGT	1200
4140	Db	CCATTATGATAAAATTAAGCATGAAAAATATTGCTGCACTACTTTTGGTGCTTAAAGTTGT	4199
1201	QY	CAC TATTCTTTGAATTAGAGTTGCTCTACAATGCACACAAATCCCGCTAAATAAATTATA	1260
4200	Db	CAC TATTCTTTGAATTAGAGTTGCTCTACAATGCACACAAATCCCGCTAAATTAATATA	4259
1261	QY	AACAAGGGTCAATTCAAATTTGAAGTAAATGTTTTATGAAGAGAGATTGAAGACAAACAG	1320
4260	Db	AACAAGGGTCAATTCAAATTTGAAGTAAATGTTTTATGAAGAGAGATTGAAGACAAACAG	4319
1321	QY	GCATAGCAAAATGACATAAGCTACCGATTACTAATCGGACACATGTA AAAACAGTTACA AAA	1380
4320	Db	GCATAGCAAAATGACATAAGCTACCGATTACTAATCGGACACATGTA AAAACAGTTACA AAA	4379
1381	QY	ATAAAGCAACTCTCTCTTCTCTACAATGAAAAGCCCTCATGTGCAGTAGAGATGCAGTT	1440
4380	Db	ATAAAGCAACTCTCTCTCTCTACAATGAAAAGCCCTCATGTGCAGTAGAGATGCAGTT	4439
1441	QY	TCATCAAGAACAAAACATCTCTTGC AAATGGGTGTAGCGGGTTCCAGATGTGAATTTGGC	1500
4440	Db	TCATCAAGAACAAAACATCTCTTGC AAATGGGTGTAGCGGGTTCCAGATGTGAATTTGGC	4499
1501	QY	AAAAACCTCATTTTAAGTAAAAAGTTTAGCAGAGCAAAAGTCGGTGCTTACGTCGCTGCTGT	1560
4500	Db	AAAAACCTCATTTTAAGTAAAAAGTTTAGCAGAGCAAAAGTCGGTGCTTACGTCGCTGCTGT	4559
1561	QY	GCGCTTGTGCGGTGGGGAGGCTCCTGCCGTGAGCTTCTTTCCCGAGCTTTGCTGCCCTGAG	1620
4560	Db	GCGCTTGTGCGGTGGGGAGGCTCCTGCCGTGAGCTTCTTTCCCGAGCTTTGCTGCCCTGAG	4619
1621	QY	AGGAACAGAGCAGACGCA CAGCGCGGAAAGCGCGCATCTAACCGCGTATCTAGGCTTTGG	1680
4620	Db	AGGAACAGAGCAGACGCA CAGCGCGGAAAGCGCGCATCTAACCGCGTATCTAGGCTTTGG	4679
1681	QY	TAACTCGGACAAGTTGCTTTTACCTGATTTGATGATACATTTTCATTAAGGTTCCAGTTA	1740
4680	Db	TAACTCGGACAAGTTGCTTTTACCTGATTTGATGATACATTTTCATTAAGGTTCCAGTTA	4739
1741	QY	TAAATATTTTGTTAATATTTATTAAGTGACTATAGAAATGCAACTCCATTTACCAAGTAACT	1800
4740	Db	TAAATATTTTGTTAATATTTATTAAGTGACTATAGAAATGCAACTCCATTTACCAAGTAACT	4799
1801	QY	TATTTTAAATATGCCWTAGTAAACATATGTAGTATATAATTTCTAGAAAACAAACATCTAATA	1860
4800	Db	TATTTTAAATATGCCWTAGTAAACATATGTAGTATATAATTTCTAGAAAACAAACATCTAATA	4859
1861	QY	AGTATATAATCTCTGCAAAATATGAGGCTTGATAATATTAGGTTGTGCAGTAGAAGCATG	1920
4860	Db	AGTATATAATCTCTGCAAAATATGAGGCTTGATAATATTAGGTTGTGCAGTAGAAGCATG	4919
1921	QY	CTAGAAGCTGTAAACAGATAACATAGAGAATAATGAGGAGTTTATGATGGAACCTTAATAT	1980
4920	Db	CTAGAAGCTGTAAACAGATAACATAGAGAATAATGAGGAGTTTATGATGGAACCTTAATAT	4979
1981	QY	ATAATGTTGCCAGGATTTTAGTTCAATATTTGTTACTGTTATCTATCTGCTGTATATGG	2040
4980	Db	ATAATGTTGCCAGGATTTTAGTTCAATATTTGTTACTGTTATCTATCTGCTGTATATGG	5039
2041	QY	AATTCCTTTTAAATCAACCGCTGAAAACGAATCAGCATTTAGCTTTGCCAGGACACCCAAA	2100
5040	Db	AATTCCTTTTAAATCAACCGCTGAAAACGAATCAGCATTTAGCTTTGCCAGGACACCCAAA	5099

Qy	2101	TAATCAGTCAATGTTAATAATATGCACAAAGTTGTGTTTTTGTTTTTGGTTTTTGTGTTTGTGTTGGTGG	2161
Db	5100	TAATCAGTCATGTTGTTAATATATGCACAAAGTTGTGTTTTTGTTTTTGGTTTTTGTGTTTGTGTTGGTGG	5159
Qy	2161	TTTTTTTGCCTTTAAAGTTGCAATGATCTTTCTGCAGGAAAATAGTCACATCATCCCACTCCACA	2220
Db	5160	TTTTTTTGCCTTTAAAGTTGCAATGATCTTTCTGCAGGAAAATAGTCACATCATCCCACTCCACA	5219
Qy	2221	TAAAGGGGTTTAGTAAGAGAAAGTCGTCTGTCTCATGATGGATAGGGGGCAAATCTTTTTTC	2280
Db	5220	TAAAGGGGTTTAGTAAGAGAAAGTCGTCTGTCTCATGATGGATAGGGGGCAAATCTTTTTTC	5279
Qy	2281	CCCTTTCTGTTAATAGTCATCACATTCTATGCCAACAGGAACGATCCATAACTTTAGT	2340
Db	5280	CCCTTTCTGTTAATAGTCATCACATTCTATGCCAACAGGAACGATCCATAACTTTAGT	5339
Qy	2341	CTTAAATGTACACATTCGATTTTGTATAAAATTAATTTTGTGTTTCCCTTTTGAGGTTTCAICG	2400
Db	5340	CTTAAATGTACACATTCGATTTTGTATAAAATTAATTTTGTGTTTCCCTTTTGAGGTTTCAICG	5399
Qy	2401	TTGCGTGTGTTTTCGCTGCACCTTTTACTTTTTCGCGTGTGGAGCTGTATTTCCCAGACAAC	2460
Db	5400	TTGCGTGTGTTTTCGCTGCACCTTTTACTTTTTCGCGTGTGGAGCTGTATTTCCCAGACAAC	5459
Qy	2461	GAAAGCGTTGGGATPACTTAAATGTAGCGACTGTCAACAGCGTGCAGGTTTTCGTGTTT	2520
Db	5460	GAAAGCGTTGGGATPACTTAAATGTAGCGACTGTCAACAGCGTGCAGGTTTTCGTGTTT	5519
Qy	2521	CTGTGTTGTGGGGTCAACCGTACAAATGGTGTGGGAATGACGATGATGTAATTTAGAA	2580
Db	5520	CTGTGTTGTGGGGTCAACCGTACAAATGGTGTGGGAATGACGATGATGTAATTTAGAA	5579
Qy	2581	TGTACCATAATTTTTTGTAATAATTTATTAATGTTTTCTAAACAAATTTATCGTATAGGTTGA	2640
Db	5580	TGTACCATAATTTTTTGTAATAATTTATTAATGTTTTCTAAACAAATTTATCGTATAGGTTGA	5639
Qy	2641	TGAAACGTCATGTTTTTGCCTAAAGACTGTAAATATTTATTTATGTTTCACATGTGTCAA	2700
Db	5640	TGAAACGTCATGTTTTTGCCTAAAGACTGTAAATATTTATTTATGTTTCACATGTGTCAA	5699
Qy	2701	AATTTCCACCACTGAAACCCCTGCACCTTAGCAACCTCATTTTTTAAAGATTAAACACAG	2760
Db	5700	AATTTCCACCACTGAAACCCCTGCACCTTAGCAACCTCATTTTTTAAAGATTAAACACAG	5759
Qy	2761	AAATAAATTTGAAAAAGGTTTTCT	2785
Db	5760	AAATAAATTTGAAAAAGGTTTTCT	5784

DEC 11 1971

RESUL 3
AX774982

LOCUS

DEFINITION

ACCESSION:

VERSION _____

KEYWORDS
CONTRACT

SOURCE ORIGIN

ORGANIC

REFERENCES

AUTHOR:

TITLE

JOURNAL:

SECRET

FEATURES

nos

ORIGIN

Query Match 100.0%; Score 2785; DB 6; Length 5784;									
Best Local Similarity 100.0%; Pred. No. 0;									
Matches 2785; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
QY	1	GAGGACTTGGACGGTGTATGACGGGGTGATTTTGTAAAGATGATTTTGTGCAATGACAAC	60						
DB	3000	GAGGACTTGGACGGTGTATGACGGGGTGATTTTGTAAAGATGATTTTGTGCAATGACAAC	3059						
QY	61	ATCCAGATATTGATGATGTGTCTGTAACCAATGCTCAGTGAGACAGACTTCAGG	120						
DB	3060	ATCCAGATATTGATGATGTGTCTGTAACCAATGCTCAGTGAGACAGACTTCAGG	3119						
QY	121	AACCTTCAGATGTGCTTGGATCCCAAGAGGACCCAAATTTGATCCCACTGGGTC	180						
DB	3120	AACCTTCAGATGTGCTTGGATCCCAAGAGGACCCAAATTTGATCCCACTGGGTC	3179						
QY	181	ATTGCCATCAAGCAAGGAGCTGGTTGACAGCCAACTCGGACCCCGGATCGCTGTA	240						
DB	3180	ATTGCCATCAAGCAAGGAGCTGGTTGACAGCCAACTCGGACCCCGGATCGCTGTA	3239						
QY	241	GGTTTTCAGAGTTTGGTCTGTGAGCTTCAGTGGCAATTCAGTAAACACTGACCGG	300						
DB	3240	GGTTTTCAGAGTTTGGTCTGTGAGCTTCAGTGGCAATTCAGTAAACACTGACCGG	3299						
QY	301	GACGACGACTATGCTGGCTTCCTCTTTGGTTACCACTCAAGCAGCGCTTCTATGTGGTG	360						
DB	3300	GACGACGACTATGCTGGCTTCCTCTTTGGTTACCACTCAAGCAGCGCTTCTATGTGGTG	3359						
QY	361	ATGTGGAGCAGGTGACGACACCTTAATGAGAGACAGCCCAACCGGGGCTTATGGCTAC	420						
DB	3360	ATGTGGAGCAGGTGACGACACCTTAATGAGAGACAGCCCAACCGGGGCTTATGGCTAC	3419						
QY	421	TCCGCGGTGCTTCAAGTGTGTAATCCACACCGGGGAGCGGGAGCACTCAGAGAAC	480						
DB	3420	TCCGCGGTGCTTCAAGTGTGTAATCCACACCGGGGAGCGGGAGCACTCAGAGAAC	3479						
QY	481	CGCTGTGGCACACGGGGAAACCGCGGGGAGGTGCGAACTTATGGCAGACCCACGG	540						
DB	3480	CGCTGTGGCACACGGGGAAACCGCGGGGAGGTGCGAACTTATGGCAGACCCACGG	3539						
QY	541	AACATTGGCTGAAGGACTACACGGCTTATAGTGGGACCTGACTCAGAGCCCAAGACC	600						
DB	3540	AACATTGGCTGAAGGACTACACGGCTTATAGTGGGACCTGACTCAGAGCCCAAGACC	3599						
QY	601	GGCTACATCAGAGTCTTAGTGATGAAGAAACAGGTTCATGGCAGACTCAGAGCTATC	660						
DB	3600	GGCTACATCAGAGTCTTAGTGATGAAGAAACAGGTTCATGGCAGACTCAGAGCTATC	3659						
QY	661	TATGACCAAACTACGCTGGCGGGGCTGGGTCTATTGTCTTCTCAAGAAATGGTC	720						
DB	3660	TATGACCAAACTACGCTGGCGGGGCTGGGTCTATTGTCTTCTCAAGAAATGGTC	3719						
QY	721	TATTTCTGAGACTCAAGTACGAATGACAGATATTAAACAAGATTTGCTGCAATTCGG	780						
DB	3720	TATTTCTGAGACTCAAGTACGAATGACAGATATTAAACAAGATTTGCTGCAATTCGG	3779						
QY	781	GCAATGCCCTGATGATGCTCCCTAGACACTCAGTTCATTTGTTGGTCTTGGGCT	840						
DB	3780	GCAATGCCCTGATGATGCTCCCTAGACACTCAGTTCATTTGTTGGTCTTGGGCT	3839						
QY	841	TCTCTCTCTAGCAGCACTTCCTGCTCCCTTGACCTTAACCTCTGATGGTTCTCACCTCTG	900						
DB	3840	TCTCTCTCTAGCAGCACTTCCTGCTCCCTTGACCTTAACCTCTGATGGTTCTCACCTCTG	3899						
QY	901	CCAGCAACCCCAACCCAGTGGCTTCAGAGTAAATCAATCAATCAATCAATCAATCAAT	960						
DB	3900	CCAGCAACCCCAACCCAGTGGCTTCAGAGTAAATCAATCAATCAATCAATCAATCAAT	3959						
QY	961	CATCTAACCCACTAGAGGAAACAGTTTGGTGATATGAGACTTTATGTGGAGTGAAA	1020						
DB	3960	CATCTAACCCACTAGAGGAAACAGTTTGGTGATATGAGACTTTATGTGGAGTGAAA	4019						
QY	1021	TTGGGCAATGCCATTAATTTGCTTTTCTTTGTTTAAAGAAATGAGCTTTACATATA	1080						
DB	4020	TTGGGCAATGCCATTAATTTGCTTTTCTTTGTTTAAAGAAATGAGCTTTACATATA	4079						
QY	1081	AAATGTAATTTACTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT	1140						
DB	4080	AAATGTAATTTACTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT	4139						
QY	1141	CCATTTATGATAAATTAAGCATGAAAAATATTTGCTGAACTACTTTTGTGCTTAAAGTTGT	1200						
DB	4140	CCATTTATGATAAATTAAGCATGAAAAATATTTGCTGAACTACTTTTGTGCTTAAAGTTGT	4199						
QY	1201	CATTTTCTTTGAATTTAGAGTGTCTTACAAATGACACAAATCCCGCTTAAATAATATA	1260						
DB	4200	CATTTTCTTTGAATTTAGAGTGTCTTACAAATGACACAAATCCCGCTTAAATAATATA	4259						
QY	1261	AACAAGGCTCAATTTCAAAATTTGAAGTAATTTTGTAGTAAGGAGAGATTAGAGACAACAG	1320						
DB	4260	AACAAGGCTCAATTTCAAAATTTGAAGTAATTTTGTAGTAAGGAGAGATTAGAGACAACAG	4319						
QY	1321	GCATGACAAATGACATTAAGCTACCGATTAATTAATCGGAAATGTAAGAGTTTACAAA	1380						
DB	4320	GCATGACAAATGACATTAAGCTACCGATTAATTAATCGGAAATGTAAGAGTTTACAAA	4379						
QY	1381	ATAAAGCACTCTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT	1440						
DB	4380	ATAAAGCACTCTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT	4439						
QY	1441	TCATCAAAAGAAACAACTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT	1500						
DB	4440	TCATCAAAAGAAACAACTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT	4499						
QY	1501	AAAACTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT	1560						
DB	4500	AAAACTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT	4559						
QY	1561	GCGGTGTGGGCTCGGGAGGCTCTCTGCTCAGCTTCTCTCCAGCTTCTGCTGCTGAG	1620						
DB	4560	GCGGTGTGGGCTCGGGAGGCTCTCTGCTCAGCTTCTCTCCAGCTTCTGCTGCTGAG	4619						
QY	1621	AGGAAACAGAGCAGACGACAGCGCGGAAAAAGGCGCATCTAACCGCTATCTAGGCTTTGG	1680						
DB	4620	AGGAAACAGAGCAGACGACAGCGCGGAAAAAGGCGCATCTAACCGCTATCTAGGCTTTGG	4679						
QY	1681	TAACTGCGGACAAAGTTGCTTTTACCTGATTTGATGATACATTTTATTAAGGTTCCAGTTA	1740						
DB	4680	TAACTGCGGACAAAGTTGCTTTTACCTGATTTGATGATACATTTTATTAAGGTTCCAGTTA	4739						
QY	1741	TAAATATTTTCTTAAATTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT	1800						
DB	4740	TAAATATTTTCTTAAATTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT	4799						
QY	1801	TATTTTAAATATGCTTAGTAAACACATATGATAGTATAATTTCTAGAAAACAAACATCTAATA	1860						
DB	4800	TATTTTAAATATGCTTAGTAAACACATATGATAGTATAATTTCTAGAAAACAAACATCTAATA	4859						
QY	1861	AGTATATATCTCTGTAATAATAGGCTTGATATATTTAGTTTCTCAGCATGAGCATG	1920						
DB	4860	AGTATATATCTCTGTAATAATAGGCTTGATATATTTAGTTTCTCAGCATGAGCATG	4919						
QY	1921	CTAGAGCTGTAAACAGAAATACATAGAGAAATATGAGGAGTTTATGATGGAACCTTAAATAT	1980						
DB	4920	CTAGAGCTGTAAACAGAAATACATAGAGAAATATGAGGAGTTTATGATGGAACCTTAAATAT	4979						
QY	1981	ATAATGTTGCCAGGATTTTGTCAATATTTTGTACTGTATCTATCTGTGTATATGG	2040						
DB	4980	ATAATGTTGCCAGGATTTTGTCAATATTTTGTACTGTATCTATCTGTGTATATGG	5039						
QY	2041	AATCTTTTAAATCAAAAGCTGAAACGAATCAGATTTAGTCTTCCAGGACACCCCAA	2100						
DB	5040	AATCTTTTAAATCAAAAGCTGAAACGAATCAGATTTAGTCTTCCAGGACACCCCAA	5099						
QY	2101	TAACTCAGTCTGTGTAATATGACAAAGTTTGTGTTTTGTTTTGTTTTGTTTTGTTTTGTTG	2160						

QY 1 GAGGACTTGACGGTGATGACCGGGTGATATTTGTTAAAGATGATTTTGACAATGACAAC 60
Db 3000 GAGGACTTGACGGTGATGACCGGGTGATATTTGTTAAAGATGATTTTGACAATGACAAC 3059
QY 61 ATCCAGATATGATGATGTGTCTGCTGAAACAAATGCCATCAGTGAGACAGACTTCAGG 120
Db 3060 ATCCAGATATGATGATGTGTCTGCTGAAACAAATGCCATCAGTGAGACAGACTTCAGG 3119
QY 121 AACTTCAGATGTCCTTGGATCCCAAGGAGACCCCAATGATCCCAACTGGTC 180
Db 3120 AACTTCAGATGTCCTTGGATCCCAAGGAGACCCCAATGATCCCAACTGGTC 3179
QY 181 ATTGCGCATCAGCGAGAGCTGGTTGACAGACCCCAATCGGACCCCGGCATCGCTGTA 240
Db 3180 ATTGCGCATCAGCGAGAGCTGGTTGACAGACCCCAATCGGACCCCGGCATCGCTGTA 3239
QY 241 GGTTTGACGAGTTGGTCTGTGGACTTCAGTGGGCACTTCTACGTAACACATGACCGG 300
Db 3240 GGTTTGACGAGTTGGTCTGTGGACTTCAGTGGGCACTTCTACGTAACACATGACCGG 3299
QY 301 GACGACACTATGCTGGCTTGTGTTTACAGTCAAGCAGCCGCTTCTATGTGGTG 360
Db 3300 GACGACACTATGCTGGCTTGTGTTTACAGTCAAGCAGCCGCTTCTATGTGGTG 3359
QY 361 ATGTGGAGCAGGTGACGACAGCTTCTGCGGAGACCCAGCCCGCGGCCTATGGCTAC 420
Db 3360 ATGTGGAGCAGGTGACGACAGCTTCTGCGGAGACCCAGCCCGCGGCCTATGGCTAC 3419
QY 421 TCCGCGGTGTCCTCAAGGTGGTGAATCCACACGGGACGGGCGAGCACTGAGGAC 480
Db 3420 TCCGCGGTGTCCTCAAGGTGGTGAATCCACACGGGACGGGCGAGCACTGAGGAC 3479
QY 481 GCGCTGTGGACACGGGAAACGCGCGGGCAGGTGCGAACCTTATGCGACGACCCGAGG 540
Db 3480 GCGCTGTGGACACGGGAAACGCGCGGGCAGGTGCGAACCTTATGCGACGACCCGAGG 3539
QY 541 AACATTGCTGGAAGGACTACACGGCTTATAGGTGGCACTGACTCAAGGCCCAAGACC 600
Db 3540 AACATTGCTGGAAGGACTACACGGCTTATAGGTGGCACTGACTCAAGGCCCAAGACC 3599
QY 601 GGCTACATCAGAGTCTTAGTGATGAAGAAACAGGTGATGCGAGACTCAGGACCTATC 660
Db 3600 GGCTACATCAGAGTCTTAGTGATGAAGAAACAGGTGATGCGAGACTCAGGACCTATC 3659
QY 661 TATGACCAAACTAGCTGGCGGGCGGTGGTCTATTTGCTTCTCTCAAGAAATGTC 720
Db 3660 TATGACCAAACTAGCTGGCGGGCGGTGGTCTATTTGCTTCTCTCAAGAAATGTC 3719
QY 721 TATTTCTCAGACCTCAAGTACGAATGACAGATATTTAAACAAGATTTGCTGCAATTCGG 780
Db 3720 TATTTCTCAGACCTCAAGTACGAATGACAGATATTTAAACAAGATTTGCTGCAATTCGG 3779
QY 781 GCAATGCCCTGTCATGCCATGGTCCCTAGACACCTCAGTTCAATTTGTTGGTCCCTGGGCT 840
Db 3780 GCAATGCCCTGTCATGCCATGGTCCCTAGACACCTCAGTTCAATTTGTTGGTCCCTGGGCT 3839
QY 841 TCTCTCTTAGCAGCACCTCCCTGACCTTAACCTGATGCTTCTCACTCTCTG 900
Db 3840 TCTCTCTTAGCAGCACCTCCCTGACCTTAACCTGATGCTTCTCACTCTCTG 3899
QY 901 CCAGCAACCCCAACCCAGTGCCTTCAGAGGATAAATCAATGGAACCTCAGAGATGAA 960
Db 3900 CCAGCAACCCCAACCCAGTGCCTTCAGAGGATAAATCAATGGAACCTCAGAGATGAA 3959
QY 961 CATCTAACCCATAGAGGAAACAGTTGGTGATATATGAGACTTTATGAGAGTGAAA 1020
Db 3960 CATCTAACCCATAGAGGAAACAGTTGGTGATATATGAGACTTTATGAGAGTGAAA 4019
QY 1021 TTGGGATGCCATTACATGCTTTTCTGTTTAAAGAAATGAGCTTTTACATATA 1080
Db 4020 TTGGGATGCCATTACATGCTTTTCTGTTTAAAGAAATGAGCTTTTACATATA 4079

QY 1081 AAATGTAATTACTTATTGTTATCTGTATATGAGTTGAAGGGAATCTGTGCATAAG 1140
Db 4080 AAATGTAATTACTTATTGTTATCTGTATATGAGTTGAAGGGAATCTGTGCATAAG 4139
QY 1141 CCATTATGATAAATTAAAGCATGAAAAATATTGCTGAACCTACTTTGGTGTCTTAAAGTTGT 1200
Db 4140 CCATTATGATAAATTAAAGCATGAAAAATATTGCTGAACCTACTTTGGTGTCTTAAAGTTGT 4199
QY 1201 CACTATTCTTGAATTAGAGTTGCTCTACAAATGACACAAATCCCGCTAAATTAATATA 1260
Db 4200 CACTATTCTTGAATTAGAGTTGCTCTACAAATGACACAAATCCCGCTAAATTAATATA 4259
QY 1261 AACAGGGTCAATTCAAATTTGAAGTAATGTTTGTAGTAAGGAGATTAGAACACACAG 1320
Db 4260 AACAGGGTCAATTCAAATTTGAAGTAATGTTTGTAGTAAGGAGATTAGAACACACAG 4319
QY 1321 GCATAGCAAAATGACATAAGCTACCGATTAACTAATCGGAACATGTAAAAAGTTACAAA 1380
Db 4320 GCATAGCAAAATGACATAAGCTACCGATTAACTAATCGGAACATGTAAAAAGTTACAAA 4379
QY 1381 ATAAACGAACTCTCTCTTGTCTTCAATGAAAGCCCTCATGTGACAGTAGAGATGCAAGTT 1440
Db 4380 ATAAACGAACTCTCTCTTGTCTTCAATGAAAGCCCTCATGTGACAGTAGAGATGCAAGTT 4439
QY 1441 TCATCAAGAACAAACATCTTGCATAATGGTGTGACCGGTTCCAGATGTGGAATTTGGC 1500
Db 4440 TCATCAAGAACAAACATCTTGCATAATGGTGTGACCGGTTCCAGATGTGGAATTTGGC 4499
QY 1501 AAAACCTCATTTAAGTAAAAAGTTAGCAGAGCAAGTCCGGTGTCTTTAGTGTCTGCTTGT 1560
Db 4500 AAAACCTCATTTAAGTAAAAAGTTAGCAGAGCAAGTCCGGTGTCTTTAGTGTCTGCTTGT 4559
QY 1561 GCCGTTGTGGCGTCGGGAGGCTCCTGCTGAGCTTCTTCCCGAGCTTTGCTGCTGAG 1620
Db 4560 GCCGTTGTGGCGTCGGGAGGCTCCTGCTGAGCTTCTTCCCGAGCTTTGCTGCTGAG 4619
QY 1621 AGGAACACAGACGACGACGCGGAAAGGCGCATCTAACGCGTATCTAGGCTTTGG 1680
Db 4620 AGGAACACAGACGACGACGCGGAAAGGCGCATCTAACGCGTATCTAGGCTTTGG 4679
QY 1681 TAACTGCGGACAAAGTTGCTTTTACCTGATTTGATGATACATTTCAATTAAGGTTCCAGTTA 1740
Db 4680 TAACTGCGGACAAAGTTGCTTTTACCTGATTTGATGATACATTTCAATTAAGGTTCCAGTTA 4739
QY 1741 TAAATATTTTGTATATTTATTAAGTACATATAGAAATGCACTCCATTTACCAGTAACT 1800
Db 4740 TAAATATTTTGTATATTTATTAAGTACATATAGAAATGCACTCCATTTACCAGTAACT 4799
QY 1801 TATTTTAATATGCTAGTAAACATATGTAGTATAATTTCTAGAAAACAAACATCTAATA 1860
Db 4800 TATTTTAATATGCTAGTAAACATATGTAGTATAATTTCTAGAAAACAAACATCTAATA 4859
QY 1861 AGTATATATCTGTGAAAAATATGAGGCTTGATAATATTTAGTGTCTCAGATGAAGCATG 1920
Db 4860 AGTATATATCTGTGAAAAATATGAGGCTTGATAATATTTAGTGTCTCAGATGAAGCATG 4919
QY 1921 CTAGAAGCTGTAAACAGATATACATAGAGAAATATAGAGGTTTATCATGGAACCTTAATAT 1980
Db 4920 CTAGAAGCTGTAAACAGATATACATAGAGAAATATAGAGGTTTATCATGGAACCTTAATAT 4979
QY 1981 ATAAATGTTGCCAGCGATTTAGTTCAAATATTTGTTACTGTTATCTATCTGCTGTATG 2040
Db 4980 ATAAATGTTGCCAGCGATTTAGTTCAAATATTTGTTACTGTTATCTATCTGCTGTATG 5039
QY 2041 AATTTCTTTTAAATCAACGCTGAAAAACGAATCAGATTTAGTTCCTGCCAGGACACCCAA 2100
Db 5040 AATTTCTTTTAAATCAACGCTGAAAAACGAATCAGATTTAGTTCCTGCCAGGACACCCAA 5099
QY 2101 TAAATCAGTCAATGTAATATGACAGGTTGTTTTGTTTTGTTTTGTTTTGTTTTGTTTTG 2160
Db 5100 TAAATCAGTCAATGTAATATGACAGGTTGTTTTGTTTTGTTTTGTTTTGTTTTGTTTTG 5159
QY 2161 TTTTTTGTCTTAAAGTTGATGATCTTTCTGCGAGGAATAGTCACTCATCCCACTCCACA 2220

2156	DB	AGGAACGAGCAGACGACAGCGCGGAAAGGCGCATCTAACGGCGTATCTTAGGCTTTGG	2211
1681	QY	TAATCTGGGGAAGAAGTTGCTTTTACCHGATTTGATGATACATTTCAATTAAGGTTCCAGTTA	1740
2216	DB	TAATCTGGGGAAGAAGTTGCTTTTACCTGATTTGATGATACATTTCAATTAAGGTTCCAGTTA	2275
1741	QY	TAATAATTTTGTAAATATTTATTTAAGTGACTATAGAATGCAACTCCATTTACAGTAACT	1800
2276	DB	TAATAATTTTGTAAATATTTATTTAAGTGACTATAGAATGCAACTCCATTTACAGTAACT	2335
1801	QY	TATTTTAAATATGCTTAGTAAACACATATGTAGTATATAATTTCTAGAAACAAACATCTAATA	1860
2336	DB	TATTTTAAATATGCTTAGTAAACACATATGTAGTATATAATTTCTAGAAACAAACATCTAATA	2395
1861	QY	AGTATATAATCTCTGTGGAATAATAGAGCCTTGATAATATTAGGTTGTCAAGATGAAGCATG	1920
2396	DB	AGTATATAATCTCTGTGGAATAATAGAGCCTTGATAATATTAGGTTGTCAAGATGAAGCATG	2455
1921	QY	CTAGAAGCTGTAAACAGATACTAGAGAATAATGAGGAGTTTATGATGAACTTT	1979
2456	DB	CTAGAAGCTGTAAACAGATACTAGAGAATAATGAGGAGTTTATGATGAACTTTTAAATA	2515
1980	QY	TATAATCTTCCAGCGATTTTGTCTCAATATTTGTATTCTGTTATCTATCTGCTGTATATG	2039
2516	DB	TATAATCTTCCAGCGATTTTGTCTCAATATTTGTATTCTGTTATCTATCTGCTGTATATG	2575
2040	QY	GAATCTCTTTTAAATCTAAACGCTGAAACGAAATGAGCATTTAGTCTGCCAGGACACCCCA	2099
2576	DB	GAATCTCTTTTAAATCTAAACGCTGAAAA - GAATCAGCATTTAGTCTGCCAGGACACCCCA	2634
2100	QY	ATAATCAGTCATGTGTAATAATGACAAAGTTTGTTTTGTTTTGTTTTGTGCTGTTG	2159
2635	DB	ATAATCAGTCATGTGTAATAATGACAAAGTTTGTTTTGTTTTGTTTTGTGCTGTTG	2694
2160	QY	----GTTTTTTTGTCTTAAAGTTGCATGATCTTTCTGCAGGAAATAGTCACTCATCCCACT	2215
2695	DB	GTTTGTCTTTTGTCTTAAAGTTGCATGATCTTTCTGCAGGAAATAGTCACTCATCCCACT	2754
2216	QY	CCACATAAGGGGTTTATGAGAGAAGTCTGCTGTGATGATGGATAGGGGCAATCT	2275
2755	DB	CCACATAAGGGGTTTATGAGAGAAGTCTGCTGTGATGATGGATAGGGGCAATCT	2814
2276	QY	TTTTTCCCTTCTGTTAATAGTCATACATTTCTATGCCAAAACAGGAACGATCCATAACT	2335
2815	DB	TTTTTCCCTTCTGTTAATAGTCATACATTTCTATGCCAAAACAGGAACGATCCATAACT	2874
2336	QY	TTAGTCTTAAATGTACACATTTGTGATAAAATTAATTTTGTGTTCCCTTTGAGGTT	2395
2875	DB	TTAGTCTTAAATGTACACATTTGTGATAAAATTAATTTTGTGTTCCCTTTGAGGTT	2934
2396	QY	GATCGTTG - - -TGTTGTTTTGCTGCACCTTTTACTTTTTTGGTGTTGGAGCTGTATCCC	2452
2935	DB	GATCGTTGTTGTTGTTTTGCTGCACCTTTTACTTTTTTGGTGTTGGAGCTGTATCCC	2994
2453	QY	GAGA - CAACGAAGCGTTGGGATCTTCATTAAGTAGCGACTGTCTCAACAGCGTGCAGGT	2511
2995	DB	GAGACCAACGAAGCGTTGGGATCTTCATTAAGTAGCGACTGTCTCAACAGCGTGCAGGT	3054
2512	QY	TTTTCTGTTTCTGTGTTGTGGGGTCAACCGTACAAATGGTGTGGGAATGACGATGTGAA	2571
3055	DB	TTTTCTGTTTCTGTGTTGTGGGGTCAACCGTACAAATGGTGTGGGAATGATGTGAA	3114
2572	QY	TATTTAGAAATGACAAATTTTTTGTGTAATTTATTTATGTTTCTTAAACAAATTTATCGT	2631
3115	DB	TATTTAGAAATGACCAATTTTTTGTGTAATTTATTTATGTTTCTTAAACAAATTTATCGT	3174
2632	QY	ATAGGTTGATGAAACGTCATGTTTTTGGCCAAAGACTGTAATATTTTATTTATGTTGTCA	2691
3175	DB	ATAGGTTGATGAAACGTCATGTTTTTGGCCAAAGACTGTAATATTTTATTTATGTTGTCA	3234
2692	QY	CATGCTCAAAATTTTCAACCACTGAAACCTCTGACTTAGCTAGAACTCTCATTTTAAAGAT	2751
3235	DB	CATGCTCAAAATTTTCAACCACTGAAACCTCTGACTTAGCTAGAACTCTCATTTTAAAGAT	3294


```

QY 2752 AACACAGGAAATAAATTTGTAAGAAAGGT 2780
Db 3295 AACACAGGAAATAAATTTGTAAGAAAGGT 3323

RESULT 8
BV177227
LOCUS sqm93505 Human DNA (Sequenom) Homo sapiens STS 10-JUN-2004
DEFINITION tagged site.
ACCESSION BV177227
VERSION BV177227.1 GI:48013259
KEYWORDS STS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 5505)
AUTHORS Nelson,R.M., Marnellos,G., Kammerer,S., Hoyal,C.R., Shi,M.M.,
Cantor,C.R. and Braun,A.
TITLE Large-Scale Validation of Single Nucleotide Polymorphisms in Gene
Regions
JOURNAL Genome Res. (2004) In press
COMMENT Contact: Andreas Braun
Pharmaceuticals division
Sequenom, Inc.
3595 John Hopkins Court, San Diego, CA 92121, USA
Tel: 18582029018
Fax: 18582029020
Email: abraun@sequenom.com
Primer A: No primer sequence submitted
Primer B: No primer sequence submitted
STS size: 5505.

FEATURES
source
1..5505
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone_lib="Human DNA (Sequenom)"
<1..>5505

STS
ORIGIN
Query Match 85.2%; Score 2372.8; DB 11; Length 5505;
Best Local Similarity 98.4%; Pred. No. 0;
Matches 2469; Conservative 1; Mismatches 29; Indels 9; Gaps 8;

QY 1 GAGGACTTGGACGGTGATGGACGGGGTGATA-TTTGTAAGATGATTTTGACAATGACAA 59
Db 3003 GAGGACTTGGACGGTGATGGACGGGGTGATA-TTTGTAAGATGATTTTGACAATGACAA 3062

QY 60 CATCCAGATATTGATGATGTGTCTGAAACAAATGCCATCAGTGAGACAGACTTCA- 118
Db 3063 CATCCAGATATTGATGATGTGTCTGAAACAAATGCCATCAGTGAGACAGACTTCCAG 3122

QY 119 GGAATCTCCAGATGTCCTTGGATCCCAAGGAGCACCAATTCATCCCAACTGGG 178
Db 3123 GGAATCTCCAGATGTCCTTGGATCCCAAGGAGCACCAATTCATCCCAACTGGG 3182

QY 179 TCATTCGCCATCAGGCAAGAGCTGTTCAGACAGCAATTCGGACCCCGGCATCGCTG 238
Db 3183 TCATTCGCCATCAGGCAAGAGCTGTTCAGACAGCAATTCGGACCCCGGCATCGCTG 3242

QY 239 TAGGTTTTGACAGTTTGGGCTGTGGACTTCAGTGCCACATCTTACGTAACACTGACC 298
Db 3243 TAGGTTTTGACAGTTTGGGCTGTGGACTTCAGTGCCACATCTTACGTAACACTGACC 3302

QY 299 GGGACGACGACTAGCTGGCTTCGTCTTGGTTACAGTCAAGCAGCGCGCTTCATGTGG 358
Db 3303 GGGACGACGACTAGCTGGCTTCGTCTTGGTTACAGTCAAGCAGCGCGCTTCATGTGG 3362

QY 359 TGATGTGAAGAGGTGACGAGACCTTACTCGGAGGACCAAGCCACGCGGCGCTATGGCT 418

```

```

Db 3363 TGATGTGAAGAGGTGACGAGACCTTCTGGAGAGCAGCCACGCGGCG--TATGGT 3420
QY 419 ACTCCGGCGTGTCCCTCAAGGTGGTGAATCCACCAACGGGGACGGCGAGCAGCTGAGGA 478
Db 3421 ACTCCGGCGTGTCCCTCAAGGTGGTGAATCCACCAACGGGGACGGCGAGCAGCTGAGGA 3480
QY 479 ACGCCGTGTGGCACACCGGGGAACACCGCGGGGAGGTGCGAACCTTTATGGCAGACCCCA 538
Db 3481 ACGCCGTGTGGCACACCGGGGAACACCGCGGGGAGGTGCGAACCTTTATGGCAGACCCCA 3540
QY 539 GGAACATTCGGCTGGAAGGACTACACGGCTATATAGTGGCACCTGACTCAGAGCCCAAGA 598
Db 3541 GGAACATTCGGCTGGAAGGACTACACGGCTATATAGTGGCACCTGACTCAGAGCCCAAGA 3600
QY 599 CCGGCTACATCAGAGCTCTTAGTGATGAAGAGAAACAGGTCTATGGCAGACTCAGGACCTA 658
Db 3601 CTGGCTACATCAGAGCTCTTAGTGATGAAGAGAAACAGGTCTATGGCAGACTCAGGACCTA 3660
QY 659 TCTATGACCAACCTACGCTGGCGGCGCTGGTCTATTGTCTTCTCTCAAGAAATGG 718
Db 3661 TCTATGACCAACCTACGCTGGCGGCGCTGGTCTATTGTCTTCTCTCAAGAAATGG 3720
QY 719 TCTATTTCTCAGACCTCAAGTACAGATATTTAAACAAGATTTGCTGCAATTC 778
Db 3721 TCTATTTCTCAGACCTCAAGTACAGATATTTAAACAAGATTTGCTGCAATTC 3780
QY 779 CGGCAATGCCCTGTGCATGCCATGCCATGCCATGCCATGCCATGCCATGCCATGCCATGCC 838
Db 3781 CGGCAATGCCCTGTGCATGCCATGCCATGCCATGCCATGCCATGCCATGCCATGCCATGCC 3840
QY 839 CTTCTCTCTTAGCAGCAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 898
Db 3841 CTTCTCTCTTAGCAGCAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3900
QY 899 TGCCAGCAACCCCAACCAAGTGCCTTCAGAGGATAAATATCAATGAACCTCAGAGATG 958
Db 3901 TGCCAGCAACCCCAACCAAGTGCCTTCAGAGGATAAATATCAATGAACCTCAGAGATG 3960
QY 959 AACATCTAACCCACTAGAGGAAACCAAGTTTGGTGATATATGAGATTTTATGTGGAGTGA 1018
Db 3961 AACATCTAACCCACTAGAGGAAACCAAGTTTGGTGATATATGAGATTTTATGTGGAGTGA 4020
QY 1019 AATTGGGCGATGCCATTACATTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCT 1078
Db 4021 AATTGGGCGATGCCATTACATTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCT 4080
QY 1079 TAAATGTAAATTACTTATTGTATTTATGTATATATGAGTTTGAAGGGAATACTGTGCATA 1138
Db 4081 TAAATGTAAATTACTTATTGTATTTATGTATATATGAGTTTGAAGGGAATACTGTGCATA 4140
QY 1139 AGCCATTATGATAAATAAGCATGAAAAATAATTCCTGAACCTACTTTTGGTGTCTTAAAGTT 1198
Db 4141 AGCCATTATGATAAATAAGCATGAAAAATAATTCCTGAACCTACTTTTGGTGTCTTAAAGTT 4200
QY 1199 GTCACTATTTCTGAAATAGAGTTGCTCTAATGACACACAATAATCCGCTCTTAAATAATTA 1258
Db 4201 GTCACTATTTCTGAAATAGAGTTGCTCTAATGACACACAATAATCCGCTCTTAAATAATTA 4260
QY 1259 TAAACAAGGGTCAATTCAAATTTCAAGTAAATGTTTATAGTAGGAGAGATTAGAGACAAC 1318
Db 4261 TAAACAAGGGTCAATTCAAATTTCAAGTAAATGTTTATAGTAGGAGAGATTAGAGACAAC 4320
QY 1319 AGGCATAGCAATAGCAATAAGCTACCGATTAACTAATCGGAACATGTAAACAGTTACAA 1378
Db 4321 AGGCATAGCAATAGCAATAAGCTACCGATTAACTAATCGGAACATGTAAACAGTTACAA 4380
QY 1379 AATAAAGCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1438
Db 4381 AATAAAGCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 4440
QY 1439 TTTTCATCAAGAGACAAACATCTCTTGCATAATGGGTGTGACCGGTTCCAGATGTGATTTG 1498

```

Db 4441 TTTCATCAAGAAACAAACATCTTTGCAAAATGGGTGTGANGCGGTTCAGATGTGGATTG 4500
Qy 1499 GCAAAACCTCATTTAAAGTAAAGGTTAGCAGAGCAAAAGTGGGTCTTTAGTCTGCTGCTT 1558
Db 4501 GCAAAACCTCATTTAAAGTAAAGGTTAGCAGAGCAAAAGTGGGTCTTTAGTCTGCTGCTT 4560
Qy 1559 GTGCCGTTGTGGCTCGGGAGGCTCTGCTGAGCTTCCTTCCCCAGCTTTGCTGCTG 1618
Db 4561 GTGCCGTTGTGGCTCGGGAGGCTCTGCTGAGCTTCCTTCCCCAGCTTTGCTGCTG 4620
Qy 1619 AGAGAAACAGAGCAGACGCAAGCGCGGAAAGCGGATCTAACGGGTATCTAGGCTTT 1678
Db 4621 AGAGAAACAGAGCAGACGCAAGCGCGGAAAGCGGATCTAACGGGTATCTAGGCTTT 4680
Qy 1679 GGTAACTGCGGACAGTTGCTTTTACCTGATTTTCATGATATATTTTCAATTAAGTTTCCAGT 1738
Db 4681 GGTAACTGCGGACAGTTGCTTTTACCTGATTTTCATGATATATTTTCAATTAAGTTTCCAGT 4740
Qy 1739 TATAAATATTTTGTAAATATTTTAAAGTACTATAGAAATCAACTCCATTTACAGTAA 1798
Db 4741 TATAAATATTTTGTAAATATTTTAAAGTACTATAGAAATCAACTCCATTTACAGTAA 4800
Qy 1799 CTTATTTTAAATATGCTTAGTAAACACATATCTAGTATATTTTCTAGAAACAAACATCTAA 1858
Db 4801 CTTATTTTAAATATGCTTAGTAAACACATATCTAGTATATTTTCTAGAAACAAACATCTAA 4860
Qy 1859 TAAGTATATATCTGTCGAAATATGAGGCTTGATAATATTAGGTGTGACAGTGAAGCA 1918
Db 4861 TAAGTATATATCTGTCGAAATATGAGGCTTGATAATATTAGGTGTGACAGTGAAGCA 4919
Qy 1919 TGCTAGAAAGCTGTAAACAGATACATAGAGAAATATGAGGAGTTTATGATGAA - CCTTAA 1977
Db 4920 TGCT - GAAGCTGT - ACCGAATACATAGAGAAATATGAGGAGTTTATGATGAA - CCTTAA 4977
Qy 1978 TATATAATGTTGCGAGCAATTTAGTCAATATTTTACTGTTATCTATCTGCTGATA 2037
Db 4978 TATATAATGTTGCGAGCAATTTAGTCAATATTTTACTGTTATCTATCTGCTGATA 5037
Qy 2038 TGGAAATCTTTTAAATCAAAAGCTGAAACAGATCAGCAATTTAGTCTTCCAGGACACACC 2097
Db 5038 TGGAAATCTTTTAAATCAAAAGCTGAAACAGATCAGCAATTTAGTCTTCCAGGACACACC 5097
Qy 2098 CAATAATCAGTCATGTAATATGACAAAGTTTGTGTTTTTGTGTTTTTGTGTTGTTGTT 2157
Db 5098 CAATAATCAGTCATGTAATATGACAAAGTTTGTGTTTTTGTGTTTTTGTGTTGTTGTT 5157
Qy 2158 TGGTTTTTTTGTCTTAAAGTTCATGATCTTTCTGAGGAAATAGTCACTCCCACTCC 2217
Db 5158 TGGTTTTTTTGTCTTAAAGTTCATGATCTTTCTGAGGAAATAGTCACTCCCACTCC 5217
Qy 2218 ACATAAGGGTTTGTAGAGAGAGTCTGCTGCTGATGATGATGAGGGGCAAAATCTTT 2277
Db 5218 ACATAAGGGTTTGTAGAGAGAGTCTGCTGCTGATGATGATGAGGGGCAAAATCTTT 5277
Qy 2278 TTCCCTTTCTGTTAATAGTCATCATTTCTATGCAAAACAGGAAAGATCCATTAATCTTT 2337
Db 5278 TTCCCTTTCTGTTAATAGTCATCATTTCTATGCAAAACAGGAAAGATCCATTAATCTTT 5337
Qy 2338 AGTCTTAATGTACACATTTGATATAAATTAATTTGTTGTTTCTTTGAGGTGCA 2397
Db 5338 AGTCTTAATGTACACATTTGATATAAATTAATTTGTTGTTTCTTTGAGGTGCA 5397
Qy 2398 TCGTTGTTGTTTGTCTGCACTTTTACTTTTTTTCGTTGTTGAGTGTATTTCCCGAGAG - 2456
Db 5398 TCGTTGTTGTTTGTCTGCACTTTTACTTTTTTTCGTTGTTGAGTGTATTTCCCGAGAG 5457
Qy 2457 CAACGAAGGCTTGGGATACCTTCAATTAATGAGGAGTGTCAACAGCG 2504
Db 5458 CAACGAAGGCTTGGGATACCTTCAATTAATGAGGAGTGTCAACAGCG 5505

LOCUS BV177983 5505 bp DNA linear STS 10-JUN-2004
DEFINITION sqm97822 Human DNA (Sequenom) Homo sapiens STS genomic, sequence tagged site.
ACCESSION BV177983
VERSION BV177983.1 GI:48014218
KEYWORDS STS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 5505)
AUTHORS Nelson,R.M., Marnellos,G., Kammerer,S., Hoyal,C.R., Shi,M.M., Cantor,C.R. and Braun,A.
TITLE Large-Scale Validation of Single Nucleotide Polymorphisms in Gene Regions
JOURNAL Genome Res. (2004) In press
COMMENT Contact: Andreas Braun
Pharmaceuticals division
Sequenom, Inc.
3595 John Hopkins Court, San Diego, CA 92121, USA
Tel: 18582029018
Fax: 18582029020
Email: abraun@sequenom.com
Primer A: No primer sequence submitted
Primer B: No primer sequence submitted
STS size: 5505.
FEATURES
source location/Qualifiers
1..5505
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone_lib="Human DNA (Sequenom)"
<1..>5505

Query Match 85.2%; Score 2372.8; DB 11; Length 5505;
Best Local Similarity 98.4%; Pred. No. 0;
Matches 2469; Conservative 1; Mismatches 29; Indels 9; Gaps 8;
Qy 1 GAGGACTTGGACGGTGTGATGGACGGGGTGATA -TTTGTAAAGATGATTTTGACAAATGACAA 59
Db 3003 GAGGACTTGGACGGTGTGATGGACGGGGTGATA -TTTGTAAAGATGATTTTGACAAATGACAA 3062
Qy 60 CATCCAGATATTTGATGATGTGTCTTGAACAAATGCCATCATGACAGACATTCAC - 118
Db 3063 CATCCAGATATTTGATGATGTGTCTTGAACAAATGCCATCATGACAGACATTCAG 3122
Qy 119 GGAACCTCCAGATGTCTCCCTTGGATCCCAAGGACCAACCAATTTGATCCCACTGGG 178
Db 3123 GGAACCTCCAGATGTCTCCCTTGGATCCCAAGGACCAACCAATTTGATCCCACTGGG 3182
Qy 179 TCATTTCCGATCAAGCAAGGAGCTGGTTTCAGACAGCAACCTCGAACCCCGCATCGCTG 238
Db 3183 TCATTTCCGATCAAGCAAGGAGCTGGTTTCAGACAGCAACCTCGAACCCCGCATCGCTG 3242
Qy 239 TAGGTTTTCACAGTTTGGTCTGTGGACTTCAGTGGACATTTCTACGTAACACTGACC 298
Db 3243 TAGGTTTTCACAGTTTGGTCTGTGGACTTCAGTGGACATTTCTACGTAACACTGACC 3302
Qy 299 GGGACGACGACTATGCTGGCTTCTGTTTGTGTTTACAGTCAAGCAGCCGCTTCTATGTGG 358
Db 3303 GGGACGACGACTATGCTGGCTTCTGTTTGTGTTTACAGTCAAGCAGCCGCTTCTATGTGG 3362
Qy 359 TGATGTGGAAGCAGGTGACGACGACTACTGGAGAGACAGCCCAAGCGGGGCTATGGCT 418
Db 3363 TGATGTGGAAGCAGGTGACGACGACTACTGGAGAGACAGCCCAAGCGGGGCTATGGCT 3420
Qy 419 ACTCCGGCTGTCCCTCAAGGTGTGAACTCCACACAGGCGGACGCGGAGACCTCAGGA 478
Db 3421 ACTCCGGCTGTCCCTCAAGGTGTGAACTCCACACAGGCGGACGCGGAGACCTCAGGA 3480
Qy 479 ACGCGCTGTGGCACACGCGGGAAACACACCGCGGGGAGGTGCGAACTTTATATGCGACGACCCCA 538

|||||
3481 AC GCGCTGTGGCACAGGGGAAACACGCGCGGGCAGGTGGCACTTTATGCGACGACCCCA 3540
QY
539 GGAACATTGCGTGGAGGACTACACGCGCTATAGTGGGACCTGACTCAGGCCCCAAGA 598
Db
3541 GGAACATTGCGTGGAGGACTACACGCGCTATAGTGGGACCTGACTCAGGCCCCAAGA 3600
QY
599 CGGCTTACATCAGAGTCTTAGTGCATGAAGGAAACACAGGTGATGGCAGAGCTCAGGACCTA 658
Db
3601 CTGGCTTACATCAGAGTCTTAGTGCATGAAGGAAACACAGGTGATGGCAGAGCTCAGGACCTA 3660
QY
659 TCTATGACCAACCTACGCTGGCGGGCGGTGGTCTATTGCTCTCTCAAGAAATGG 718
Db
3661 TCTATGACCAACCTACGCTGGCGGGCGGTGGTCTATTGCTCTCTCAAGAAATGG 3720
QY
719 TCTATTCTCAGACCTCAAGTACGATGAGAGATATTTAAACAGATTTGCTGCATTTTC 778
Db
3721 TCTATTCTCAGACCTCAAGTACGATGAGAGATATTTAAACAGATTTGCTGCATTTTC 3780
QY
779 CGGCAATGCCCTGTGCATGCCATGCTGCTTCCCTAGACACCTCAGATTCTATTGCTGCTTGGCG 838
Db
3781 CGGCAATGCCCTGTGCATGCCATGCTGCTTCCCTAGACACCTCAGTTCTATTGCTGCTTGGCG 3840
QY
839 CTTCTCTCTAGCAGACCTCTCTGCTCCCTTGACCTTAACTCTGATGTTCTTCACTCC 898
Db
3841 CTTCTCTCTAGCAGACCTCTCTGCTCCCTTGACCTTAACTCTGATGTTCTTCACTCC 3900
QY
899 TGCCAGCAACCCCAACCAAGTCCCTCAGAGGATAATATCAATGGAACCTCAGAGATG 958
Db
3901 TGCCAGCAACCCCAACCAAGTCCCTCAGAGGATAATATCAATGGAACCTCAGAGATG 3960
QY
959 AACATCTAAACCCACATAGAGAAACCAAGTTGGTGATATATGAGACTTTATGTTGGAGTGAA 1018
Db
3961 AACATCTAAACCCACATAGAGAAACCAAGTTGGTGATATATGAGACTTTATGTTGGAGTGAA 4020
QY
1019 AATTGGGCATGCCATTACATGCTTTTCTGTTGTTTAAAGAAATGACGTTTACATA 1078
Db
4021 AATTGGGCATGCCATTACATGCTTTTCTGTTGTTTAAAGAAATGACGTTTACATA 4080
QY
1079 TAAATGTAAATTTACTATTGTTATTTATGTTATGTTATGAGTTGAAGGATATCTGTCATA 1138
Db
4081 TAAATGTAAATTTACTATTGTTATTTATGTTATGTTATGAGTTGAAGGATATCTGTCATA 4140
QY
1139 AGCCATTATGATAAATTAAGCATGAAAAATATGCTGAACTACTTTTGGTGTCTTAAAGTT 1198
Db
4141 AGCCATTATGATAAATTAAGCATGAAAAATATGCTGAACTACTTTTGGTGTCTTAAAGTT 4200
QY
1199 GTCACATTTCTTGAATTAGAGTTGCTCTCAATGACACACAAATCCCGCTAAATAATTA 1258
Db
4201 GTCACATTTCTTGAATTAGAGTTGCTCTCAATGACACACAAATCCCGCTAAATAATTA 4260
QY
1259 TAAACAAGGTCATTTCAAATTTGAAGTAATGTTTATGTAAGGAGAGATTAGAAGCAAC 1318
Db
4261 TAAACAAGGTCATTTCAAATTTGAAGTAATGTTTATGTAAGGAGAGATTAGAAGCAAC 4320
QY
1319 AGGCATAGCAATGACATAAGCTACCGATTAACTAATCGGAACATGTAACACAGTTACAA 1378
Db
4321 AGGCATAGCAATGACATAAGCTACCGATTAACTAATCGGAACATGTAACACAGTTACAA 4380
QY
1379 AAATAAGCAACTCTCTCTTCTGCTTACAAATGAAAGCCCTCATGTGCAAGTACAGATG 1438
Db
4381 AAATAAGCAACTCTCTCTTCTGCTTACAAATGAAAGCCCTCATGTGCAAGTACAGATG 4440
QY
1439 TTTTCATCAAGAAACAAACATCTCTCAATGAGTGGTGACGCGGTTCCAGATGATGATTG 1498
Db
4441 TTTTCATCAAGAAACAAACATCTCTCAATGAGTGGTGACGCGGTTCCAGATGATGATTG 4500
QY
1499 GCAAAACCTCATTTAAGTAAAGGTTAGCAGAGCAAGTGGCGTCTTTAGCTGCTGCTT 1558
Db
4501 GCAAAACCTCATTTAAGTAAAGGTTAGCAGAGCAAGTGGCGTCTTTAGCTGCTGCTT 4560
QY
1559 GTGCGGTTGTGGCGTGGGGAGGCTCTGCTGAGCTTCTTCCCAAGCTTTGCTGCTG 1618
|||||

Db
4561 GTGCGCTGTGGNGTCGGGAGGCTCTCGCTGAGCTTCTTCCCGAGCTTTGCTGCCTG 4620
QY
1619 AGAGGAACAGAGCAGACGACAGCGCGGAAAGGCGCATCTAAACGCGTATCTAGGCTTT 1678
Db
4621 AGAGGAACAGAGCAGACGACAGCGCGGAAAGGCGCATCTAAACGCGTATCTAGGCTTT 4680
QY
1679 GGTAACTCGGGAACAAGTTGCTTTTACCTGATTGATGATACATTTCAATTAAGGTTCCAGT 1738
Db
4681 GGTAACTCGGGAACAAGTTGCTTTTACCTGATTGATGATACATTTCAATTAAGGTTCCAGT 4740
QY
1739 TATAAATATTTTGTAAATATTTAATAGTGACTATAGAAATGCAACTCCATTTTACCAGTAA 1798
Db
4741 TATAAATATTTTGTAAATATTTAATAGTGACTATAGAAATGCAACTCCATTTTACCAGTAA 4800
QY
1799 CTTATTTTAAATATGCTCTAGTAAACATATGATGATATAATTTCTAGAAAACAAACATCTAA 1858
Db
4801 CTTATTTTAAATATGCTCTAGTAAACATATGATGATATAATTTCTAGAAAACAAACATCTAA 4860
QY
1859 TAAAGTATATAAATCCCTGTGAAAAATATGAGGCTTTGATAATATAGGTTGTCCAGATGAAGCA 1918
Db
4861 TAAAGTATATAAATCCCTGTGAAAAATATGAGGCTTTGATAATATAGGTTGTCCAGATG-AGCA 4919
QY
1919 TCTAGAAGCTGTAACAGAAATACATAGAGAAATAATGAGGAGTTTATGATGGAA-CCTTAA 1977
Db
4920 TGCT-GAAGCTGT-ACCGAATACATAGAGAAATAATGAGAGTTGATGATGAAACCTTAA 4977
QY
1978 TATATAATGTTGCCAGCGATTTTAGTTCAATATTTGTTACTGTTATCTATCTGCTGTATA 2037
Db
4978 TATATAATGTTGCCAGCGATTTTAGTTCAATATTTGTTACTGTTATCTATCTGCTGTATA 5037
QY
2038 TCGAATTTCTTTTAAATTTCAAAACGCTGAAACGAATCAGCAATTTAGTCTTCCAGGACACC 2097
Db
5038 TCGAATTTCTTTTAAATTTCAAAACGCTGAAACGAATCAGCAATTTAGTCTTCCAGGACACC 5097
QY
2098 CAATAATCAGTCATGTAATAATGACAAAGTTTGTGTTTTTGTGTTTTTGTGTTGTTGTT 2157
Db
5098 CAATAATCAGTCATGTAATAATGACAAAGTTTGTGTTTTTGTGTTTTTGTGTTGTTGTT 5157
QY
2158 TGGTTTTTGTGTTTAAAGTTGATGATCTTCTGCGAGGAAATAGTCACTCATCCCACTCC 2217
Db
5158 TGGTTTTTGTGTTTAAAGTTGATGATCTTCTGCGAGGAAATAGTCACTCATCCCACTCC 5217
QY
2218 ACATAAGGGTTTGTAAAGAGAGTCTGCTGCTGATGATGATGAGGCGCAAAATCTTT 2277
Db
5218 ACATAAGGGTTTGTAAAGAGAGGNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 5277
QY
2278 TTCCCTTTTCTGTTTAAATAGTCACTACATTTCTATGCCAAACAGGAACGATCCATAACTTT 2337
Db
5278 TTCCCTTTTCTGTTTAAATAGTCACTACATTTCTATGCCAAACAGGAACGATCCATAACTTT 5337
QY
2338 AGTCTTAATGATACATTTGCACTTTTGTATAAATTAATTTTGTGTTTCTTTGAGTTGA 2397
Db
5338 AGTCTTAATGATACATTTGCACTTTTGTATAAATTAATTTTGTGTTTCTTTGAGTTGA 5397
QY
2398 TCGTTGTGTTGTTTGTGTCACCTTTTACTTTTTTGTGCGTGTGGAGCTGTATTTCCCGAG- 2456
Db
5398 TCGTTGTGTTGTTTGTGTCACCTTTTACTTTTTTGTGCGTGTGGAGCTGTATTTCCCGAG- 5457
QY
2457 CAACGAAGCGTTGGGATACCTTCAATTAATGATGAGGACTGTCAACAGCG 2504
Db
5458 CAACGAAGCGTTGGGATACCTTCAATTAATGATGAGGACTGTCAACAGCG 5505

RESULT 10
BV178635
LOCUS
DEFINITION
tagged site.
ACCESSION
BV178635
VERSION
BV178635.1
KEYWORDS
STs.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens

BV178635
5505 bp
DNA
linear
STS 10-JUN-2004
Human DNA (Sequenc)
Homo sapiens STS genomic, sequence

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
AUTHORS

1 (bases 1 to 5505)
Nelson, R.M., Marnellos, G., Kammerer, S., Hoyal, C.R., Shi, M.M.,
Cantor, C.R. and Braun, A.

Large-Scale Validation of Single Nucleotide Polymorphisms in Gene
Regions

JOURNAL
COMMENT

Genome Res. (2004) In press

Contact: Andreas Braun
Pharmaceuticals division
Sequenom, Inc.

3595 John Hopkins Court, San Diego, CA 92121, USA

Tel: 18582023018

Fax: 18582023020

Email: abraun@sequenom.com

Primer A: No primer sequence submitted

Primer B: No primer sequence submitted

STS size: 5505.

FEATURES
Source

Location/Qualifiers

1..5505

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

/clone_lib="Human DNA (Sequenom)"

<1..5505

ORIGIN

Query Match 85.2%; Score 2372.8; DB 11; Length 5505;
Best Local Similarity 98.4%; Pred. No. 0;
Matches 2469; Conservative 1; Mismatches 29; Indels 9; Gaps 8;

QY 1 GAGGACTTGGACGGGTGATGGCGGGTGTGATA-TTTGTAAGATGATTTTGACAATGACAA 59
DB 3003 GAGGACTTGGACGGGTGATGGCGGGTGTGATA-TTTGTAAGATGATTTTGACAATGACAA 3062

QY 60 CATCCAGATATGATGATGTGTGCTGAAACAAATGCCATCAGTGAGACAGACTTCA- 118
DB 3063 CATCCAGATATGATGATGTGTGCTGAAACAAATGCCATCAGTGAGACAGACTTCA 3122

QY 119 GGAATTCAGATGTCCTTGGATCCCAAGGAGACACCCAAATGATCCCACTGGG 178
DB 3123 GGAATTCAGATGTCCTTGGATCCCAAGGAGACACCCAAATGATCCCACTGGG 3182

QY 179 TCATTTCGCATCAAGGAGAGCTGTTTCAGACAGCCAACTCGACCCCGGCATCGTG 238
DB 3183 TCATTTCGCATCAAGGAGAGCTGTTTCAGACAGCCAACTCGACCCCGGCATCGTG 3242

QY 239 TAGGTTTTCAGAGTTTGGGCTGTGGACTTCAGTGGCAATTCACGTAACACTGACC 298
DB 3243 TAGGTTTTCAGAGTTTGGGCTGTGGACTTCAGTGGCAATTCACGTAACACTGACC 3302

QY 299 GGGACAGACATGCTGCTTGGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGT 358
DB 3303 GGGACAGACATGCTGCTTGGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGT 3362

QY 359 TGATGTGAAGCAGGTGACGACAGCTACTGAGGAGACACAGCCAGCGGGCTATGGCT 418
DB 3363 TGATGTGAAGCAGGTGACGACAGCTACTGAGGAGACACAGCCAGCGGGCTATGGCT 3420

QY 419 ACTCCGGGCTGCTCCTCAAGTGGTGAATCCACACAGGGGACGGGCGAGCAGCTGAGGA 478
DB 3421 ACTCCGGGCTGCTCCTCAAGTGGTGAATCCACACAGGGGACGGGCGAGCAGCTGAGGA 3480

QY 479 AGCGGTGTGGACACAGGGAGACACCGCGGGCAGGTGCGAACCTTATGGCAGCCCA 538
DB 3481 AGCGGTGTGGACACAGGGAGACACCGCGGGCAGGTGCGAACCTTATGGCAGCCCA 3540

QY 539 GGAACATTTGGCTGGAAGGACTACACGGCTATAGTGGCAGCTGACTCAGAGCCCAAGA 598
DB 3541 GGAACATTTGGCTGGAAGGACTACACGGCTATAGTGGCAGCTGACTCAGAGCCCAAGA 3600

QY 599 CCGGCTACATCAGAGCTTTAGTGCATGAAGGAAAACAGGTGATGGCAGACTCAGGACCTA 658

DB 3601 CTGGCTACATCAGAGTCTTTAGTGCATGAAGAAAACAGGTCATGGCAGACTCAGGACCTA 3660

QY 659 TCTATGACCAAACTACGCTGGCGGGCTGGGTCTATTTGTTCTTCTCTCAAGAAATGG 718

DB 3661 TCTATGACCAAACTACGCTGGCGGGCTGGGTCTATTTGTTCTTCTCTCAAGAAATGG 3720

QY 719 TCTATTTCTCAGACCTCAAGTACGAATGACAGATATTTAAACAAGATTTGCTGCAATTC 778

DB 3721 TCTATTTCTCAGACCTCAAGTACGAATGACAGATATTTAAACAAGATTTGCTGCAATTC 3780

QY 779 CGCAATGCTCTGTGATGCCATGGTCCCTAGACACCTCAGTTCAATGTTGCTTGGTGG 838

DB 3781 CGCAATGCTCTGTGATGCCATGGTCCCTAGACACCTCAGTTCAATGTTGCTTGGTGG 3840

QY 839 CTCTCTCTCTAGCAGACCTCTCTGCTTGAACCTTAACCTCTGATGTTCTTCACTCC 898

DB 3841 CTCTCTCTCTAGCAGACCTCTCTGCTTGAACCTTAACCTCTGATGTTCTTCACTCC 3900

QY 899 TGCCAGCAACCCCAACCCAAAGTGCCTTTCAGAGATTAATATCAATGAACTCAGAGATG 958

DB 3901 TGCCAGCAACCCCAACCCAAAGTGCCTTTCAGAGATTAATATCAATGAACTCAGAGATG 3960

QY 959 AACATCTAACCCACTAGAGGAAACAGTTTGGTGATATATGAGACTTTATGAGAGTCAA 1018

DB 3961 AACATCTAACCCACTAGAGGAAACAGTTTGGTGATATATGAGACTTTATGAGAGTCAA 4020

QY 1019 AATTGGGATGCCATTACATTTCTTTCTTTGTTTGTAAAGAAATGACGTTTACATA 1078

DB 4021 AATTGGGATGCCATTACATTTCTTTGTTTGTAAAGAAATGACGTTTACATA 4080

QY 1079 TAAATGTAATTAATTAATTTATGTTATGTTATGAGGATTAAGGAAATTAATGTTGATA 1138

DB 4081 TAAATGTAATTAATTAATTTATGTTATGTTATGAGGATTAAGGAAATTAATGTTGATA 4140

QY 1139 AGCCATTAATTAATTAAGCATGAATAATTTGCTGAACCTTCTTCTGCTGCTTAAAGTT 1198

DB 4141 AGCCATTAATTAATTAAGCATGAATAATTTGCTGAACCTTCTTCTGCTGCTTAAAGTT 4200

QY 1199 GTCACTATTTCTGTAATTAAGTTGCTCTACAATGACACACAAATCCCGCTTAATAATTA 1258

DB 4201 GTCACTATTTCTGTAATTAAGTTGCTCTACAATGACACACAAATCCCGCTTAATAATTA 4260

QY 1259 TAAACAAGGTCATTTCAAAATTTGAAGTAATTTAGTAAAGAGATTAAGAGACAAAC 1318

DB 4261 TAAACAAGGTCATTTCAAAATTTGAAGTAATTTAGTAAAGAGATTAAGAGACAAAC 4320

QY 1319 AGGCATAGCAATGACATAGCTACCGATTAACTAATCGGACATGTAACAGTTACAA 1378

DB 4321 AGGCATAGCAATGACATAGCTACCGATTAACTAATCGGACATGTAACAGTTACAA 4380

QY 1379 AAATAAAGCACTCTCTCTTGTCTACAATGAAGCCCTCATGTGCAGTAGAGATGCAG 1438

DB 4381 AAATAAAGCACTCTCTCTTGTCTACAATGAAGCCCTCATGTGCAGTAGAGATGCAG 4440

QY 1439 TTTTCATCAAGAACAAACATCTCTGCAATGGGTGTGACGGTTCAGAGATGGAATTTG 1498

DB 4441 TTTTCATCAAGAACAAACATCTCTGCAATGGGTGTGACGGTTCAGAGATGGAATTTG 4500

QY 1499 GCATAACCTCATTTAAGTAAAGGTTAGCAGAGCAAGTGGCGTGTCTTTAGCTGCTGCTT 1558

DB 4501 GCATAACCTCATTTAAGTAAAGGTTAGCAGAGCAAGTGGCGTGTCTTTAGCTGCTGCTT 4560

QY 1559 GTGCGGTTGTGGCGTGGGGAGGCTCTGCTGCTGAGCTTCTTCCCAAGCTTTGCTGCC 1618

DB 4561 GTGCGGTTGTGGCGTGGGGAGGCTCTGCTGCTGAGCTTCTTCCCAAGCTTTGCTGCC 4620

QY 1619 AGAGAAACAGAGCAGACGACAGCGCGGAAAGAGCGCATCTAAACGCTATCTAGGCTTT 1678

DB 4621 AGAGAAACAGAGCAGACGACAGCGCGGAAAGAGCGCATCTAAACGCTATCTAGGCTTT 4680

QY 1679 GGTAACTGCGGCAAGTTGCTTTTAACTGATTTGATACATTTTCAATTAAGGTTCCAGT 1738

JOURNAL

Genome Res. (2004) In press

Genome Res. (2004) In press

Qy 1470 GGTGTGACGCGGTTCCAGATGTGGATTGGCAAAACCTCATTTAAGTAAAGGTTAGCAG 1522

Db 28091 GGTGTGACGCGGTTCCAGATGTGGATTGGCAAAACCTCATTTAAGTAAAGGTTAGCAG 28033

The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information on the WORMPEP database can be found at

Qy	1530	AGCAAAGTCGGGTGCTTTAGCTGCTGCTGTGTGCGGTGTGGCGTCCGGAGAGCTCTCTGCC	1589
Db	28031	AGCAAAGTCGGGTGCTTTAGCTGCTGCTGTGTGCGGTGTGGCGTCCGGAGAGCTCTCTGCC	27972
Qy	1590	TGAGCTTCCTTCCCAGCTTTGCTGCTCAGAGGACACAGACAGACGACAGGCCGGA	1649
Db	27971	TGAGCTTCCTTCCCAGCTTTGCTGCTCAGAGGAAACAGACAGACGACAGGCCGGA	27912
Qy	1650	AAGCGCATCTAAACGGTATCTAGGCTTTGGTAACTGCGGACAAAGTGTGCTTTTACCTCAT	1709
Db	27911	AAGCGCATCTAAACGGTATCTAGGCTTTGGTAACTGCGGACAAAGTGTGCTTTTACCTCAT	27852
Qy	1710	TTGATGATACATTTCAITTAAGGTTCCAGTTATAAATATATTTTGGTAAATTTATTTAAAGTGA	1769
Db	27851	TTGATGATACATTTCAITTAAGGTTCCAGTTATAAATATATTTTGGTAAATTTATTTAAAGTGA	27792
Qy	1770	CTATAGAATGCAACTCCATTTACCACTTAATTTTAAATATGCTAGTAACACATATG	1829
Db	27791	CTATAGAATGCAACTCCATTTACCACTTAATTTTAAATATGCTAGTAACACATATG	27732
Qy	1830	TAGTATAAATTTCTAGAAACAAACATCTAATAAGTATATAAATCCCTGTGAAATATGAGGCT	1889
Db	27731	TAGTATAAATTTCTAGAAACAAACATCTAATAAGTATATAAATCCCTGTGAAATATGAGGCT	27672
Qy	1890	TGATAATATAGGTTGTGAGATGAAGCATGCTAGAGCTGTAACAGAAATACATAGAGAA	1949
Db	27671	TGATAATATAGGTTGTGAGATGAAGCATGCTAGAGCTGTAACAGAAATACATAGAGAA	27612
Qy	1950	TAATGAGGAGTTTATCATGGAACCTT - AATATATAATGTTGCCAGCGATTTTAGTTCAT	2008
Db	27611	TAATGAGGAGTTTATCATGGAACCTTAAATATAATATGTTGCCAGCGATTTTAGTTCAT	27552
Qy	2009	ATTGTTACTGTTATCTATCTGTTATATGAAATCTTTTAAATCAAACGCTCAAACG	2068
Db	27551	ATTGTTACTGTTATCTATCTGTTATATGAAATCTTTTAAATCAAACGCTCAAACG	27493
Qy	2069	AATCAGCATTTAGTCTTGCCAGGCACCCCAATAATCAGTCATGTATAATGACAAAGT	2128
Db	27492	AATCAGCATTTAGTCTTGCCAGGCACCCCAATAATCAGTCATGTATAATGACAAAGT	27433
Qy	2129	TGTGTTTTGTTTTGTTTTTTTTTTGTTGGTTG --- GTTTTTTGTGTTTAAAGTTCGATGAT	2184
Db	27432	TGTGTTTTGTTTTGTTTTTTTTTTGTTGGTTGTTTTTGTGCTTTAAAGTTCGATGAT	27373
Qy	2185	CTTCTGACGAAATAGTCACATCCCACTCCACATAAGGGGTTTAGTAAGAGAAAGTCT	2244
Db	27372	CTTCTGACGAAATAGTCACATCCCACTCCACATAAGGGGTTTAGTAAGAGAAAGTCT	27313
Qy	2245	GTCGTCTGATGATGATAGGGGGGAAATCTTTTCCCTTTCTGTGTTAAATGATCATCAC	2304
Db	27312	GTCGTCTGATGATGATAGGGGGGAAATCTTTTCCCTTTCTGTGTTAAATGATCATCAC	27253
Qy	2305	TTTCTATGCAAAACAGGAACGATCCATAACCTTTAGTCTTAATGTACACATTCGATTTTGA	2364
Db	27252	TTTCTATGCAAAACAGGAACGATCCATAACCTTTAGTCTTAATGTACACATTCGATTTTGA	27193
Qy	2365	TAAAAATTAATTTGTTGTTTCTTTGAGGTTGATCGTTG --- TGTGTGTTTTGCTGCACCTT	2421
Db	27192	TAAAAATTAATTTGTTGTTTCTTTGAGGTTGATCGTTGTTGTTGTTGCTGCACCTT	27133
Qy	2422	TTTATCTTTTTCGGTGTGAGCTGATTTCCGAGA - CAACGAACGGTGGGATATCTTCAT	2480
Db	27132	TTTATCTTTTTCGGTGTGAGCTGATTTCCGAGACCAACGAACGGTTGGGATACTTCAT	27073
Qy	2481	TAAATGTAGCGACTGTCACACGGTGCAGGTTTCTGTTCTGTGTTGTGGGTCACACCG	2540
Db	27072	TAAATGTAGCGACTGTCACACGGTGCAGGTTTCTGTTCTGTGTTGTGGGTCACACCG	27013
Qy	2541	TACAATGGTGTGGGAATACCGATGATGGAATATTTAGAATGTACCATATTTTTGTGAAA	2600
Db	27012	TACAATGGTGTGGGAATACCGATGATGGAATATTTAGAATGTACCATATTTTTGTGAAA	26953
Qy	2601	TTATTTATGTTTTTCTAAACAAATTTATCGTATAGGTTGATGAACCGCTCATGTGTTTGC	2660

[illegible]

QY 361 ATGTGGAGCAGGTGACGACAGACCTACTGGGAGGACACAGCCACGCGGCGCTATGGCTAC 420
Db 1445 ATGTGGAGCAGGTGACGACAGACCTACTGGGAGGACACAGCCACGCGGCGCTATGGCTAC 1504
QY 421 TCCGCGGTGTCCTTCAAGGTGGTGAATCCACACCGGGGACGGCGGACGACCTCAGGAAC 480
Db 1505 TCCGCGGTGTCCTTCAAGGTGGTGAATCCACACCGGGGACGGCGGACGACCTCAGGAAC 1564
QY 481 GCGTGTGGCACAACGGGGAACACGCGGGGAGGTGGAACTTATGGCAGACCCGAGG 540
Db 1565 GCGTGTGGCACAACGGGGAACACGCGGGGAGGTGGAACTTATGGCAGACCCGAGG 1624
QY 541 AACATTGGCTGGAAGGACTACACGCGCTATAGGTGGACCTGACTCAACAGCCCAAGACC 600
Db 1625 AACATTGGCTGGAAGGACTACACGCGCTATAGGTGGACCTGACTCAACAGCCCAAGACC 1684
QY 601 GGCTACATCAGAGTCTTAGTGCATGAAGGAAAAACAGGTTCATGGCAGACTCAGGACCTATC 660
Db 1685 GGCTACATCAGAGTCTTAGTGCATGAAGGAAAAACAGGTTCATGGCAGACTCAGGACCTATC 1744
QY 720 TATGACCAAACTAGCTGGGGGGGCGGTCTATTGTCTCTCTCAAGAAATGTC 720
Db 1745 TATGACCAAACTAGCTGGGGGGGCGGTCTATTGTCTCTCTCAAGAAATGTC 1804
QY 721 TATTTCTCAGACCTCAAGTACGAATGCAGAGATATTTAAACAGAAATTTGCTGCAATTCGG 780
Db 1805 TATTTCTCAGACCTCAAGTACGAATGCAGAGATATTTAAACAGAAATTTGCTGCAATTCGG 1864
QY 781 GCAATGCCCTGTGATGTCATGCTCCCTAGACACCTCAGTTCATTTGGTCTCTCGGCT 840
Db 1865 GCAATGCCCTGTGATGTCATGCTCCCTAGACACCTCAGTTCATTTGGTCTCTCGGCT 1924
QY 841 TCTCTCTAGCAGACCTCTCTGCTCCCTGACCTTAACTCTGATGTTCTTCACTCTG 900
Db 1925 TCTCTCTAGCAGACCTCTCTGCTCCCTGACCTTAACTCTGATGTTCTTCACTCTG 1984
QY 901 CCAGCAACCCCAACCCCAAGTCCCTTCAGAGGATAAATATCAATGGAACCTCAGAGATGAA 960
Db 1985 CCAGCAACCCCAACCCCAAGTCCCTTCAGAGGATAAATATCAATGGAACCTCAGAGATGAA 2044
QY 961 CATCTAACCCCTAGAGGAAACCAAGTTGGTGATATATGAGACTTTATGTGGAGTGAATA 1020
Db 2045 CATCTAACCCCTAGAGGAAACCAAGTTGGTGATATATGAGACTTTATGTGGAGTGAATA 2104
QY 1021 TTGGGCATGCCATTAATGCTTTCTGTTTGTAAAGAGATGAGTTTACATATA 1080
Db 2105 TTGGGCATGCCATTAATGCTTTCTGTTTGTAAAGAGATGAGTTTACATATA 2164
QY 1081 AAATGTAATTACTTATTTATTTATGTTATGTTGAGTTGAAGGGAATCTGTGCATAAG 1140
Db 2165 AAATGTAATTACTTATTTATTTATGTTATGTTGAGTTGAAGGGAATCTGTGCATAAG 2224
QY 1141 CCATTATGATAAATTAAGCATGAAATAATTCGTAACCTATTTTGGTCTTAAAGTTGT 1200
Db 2225 CCATTATGATAAATTAAGCATGAAATAATTCGTAACCTATTTTGGTCTTAAAGTTGT 2284
QY 1201 CACTATTCTTGAATTAGAGTCTCTACAAATGACACACAAATCCCGCTAAATAAATTATA 1260
Db 2285 CACTATTCTTGAATTAGAGTCTCTACAAATGACACACAAATCCCGCTAAATAAATTATA 2344
QY 1261 AACAAAGGCTCAATTTCAAAATTTGAAAGTAATGTTTATGTAAGGAGAGATTAGAGACACAG 1320
Db 2345 AACAAAGGCTCAATTTCAAAATTTGAAAGTAATGTTTATGTAAGGAGAGATTAGAGACACAG 2404
QY 1321 GCATAGCAATGATCAATGATCCGATTAATCTAATCGGAAACATGTAAGACAGTTACAAA 1380
Db 2405 GCATAGCAATGATCAATGATCCGATTAATCTAATCGGAAACATGTAAGACAGTTACAAA 2464
QY 1381 ATAAACGAACTCTCTCTCTCTCAATGAAGCCCTCATGTGCAAGTATGAGATGAGTT 1440
Db 2465 ATAAACGAACTCTCTCTCTCTCAATGAAGCCCTCATGTGCAAGTATGAGATGAGTT 2524

QY 1441 TCATCAAGAACAAACATCTCTTGCATATGGTGTGACGCGGTTCCAGATGTGATTTGGC 1500
Db 2525 TCATCAAGAACAAACATCTCTTGCATATGGTGTGACGCGGTTCCAGATGTGATTTGGC 2584
QY 1501 ABAACCTCATTTAAGTAAAGGTTAGCAGAGCAAAAGTGGGTGCTTTAGCTGCTGCTTGT 1560
Db 2585 ABAACCTCATTTAAGTAAAGGTTAGCAGAGCAAAAGTGGGTGCTTTAGCTGCTGCTTGT 2644
QY 1561 GCCCTGTGGCGTCCGGGAGGCTCCTGCTAGCTTCTCTCCAGCTTTGCTCCCTGAG 1620
Db 2645 GCCCTGTGGCGTCCGGGAGGCTCCTGCTAGCTTCTCTCCAGCTTTGCTCCCTGAG 2704
QY 1621 AGGAACACAGACGACGACGACGCGGCAAAAGGCGCATCTAACCGTATCTAGGCTTGG 1680
Db 2705 AGGAACACAGACGACGACGACGCGGCAAAAGGCGCATCTAACCGTATCTAGGCTTGG 2764
QY 1681 TAACCTGCGGCAAG 1694
Db 2765 TAACCTGCGGCAAG 2778

RESULT 15
BV177085/c
LOCUS BV177085 Human DNA (Sequenom) Homo sapiens STS 10-JUN-2004
DEFINITION tagged site.
ACCESSION BV177085
VERSION BV177085.1 GI:48013080
KEYWORDS STS,
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 921)
AUTHORS Nelson, R.M., Marnellos, G., Kammerer, S., Hoyal, C.R., Shi, M.M.,
Cantor, C.R. and Braun, A.
TITLE Large-Scale Validation of Single Nucleotide Polymorphisms in Gene
Regions
JOURNAL Genome Res. (2004) In press
COMMENT
Contact: Andreas Braun
Pharmaceuticals division
Sequenom, Inc.
3595 John Hopkins Court, San Diego, CA 92121, USA
Tel: 18582029018
Fax: 18582029020
Email: abraun@sequenom.com
Primer A: No primer sequence submitted
Primer B: No primer sequence submitted
STS size: 921.
Location/Qualifiers
1..921
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone_lib="Human DNA (Sequenom)"
<1..>921

FEATURES
source

STS
ORIGIN

Query Match 29.0%; Score 809; DB 11; Length 921;
Best Local Similarity 98.2%; Pred. No. 6.1e-168;
Matches 869; Conservative 1; Mismatches 5; Indels 10; Gaps 5;

QY 1910 GATGAAGCATGCTAGAAGCTGTAAACAGATAACATAGAGAAATATGAGGATTTATGATCG 1969
Db 915 GATGAAGCATGCTAGAAGCTGTAAACAGATAACATAGAGAAATATGAGGATTTATGATCG 856
QY 1970 AACCTTTAATATATATGTTGCCAGGATTTTACTTCATATTTTACTGTTATCTATC 2028
Db 855 AACCTTTAATATATATGTTGCCAGGATTTTACTTCATATTTTACTGTTATCTATC 796
QY 2029 TCGTGTATATGAAATCTTTTAAATCAAAACGCTGAAACGAATCAGCAATTTAGTCTTGGC 2088

This Page Blank (uspto)